ESTCP
Cost and Performance Report

(MM-0034)

Advanced UXO Detection/Discrimination Technology Demonstration

November 2005

ENVIRONMENTAL SECURITY
TECHNOLOGY CERTIFICATION PROGRAM

U.S. Department of Defense

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **NOV 2005**	2. REPORT TYPE	3. DATES COVERED **00-00-2005 to 00-00-2005**

4. TITLE AND SUBTITLE **Advanced UXO Detection/Discrimination Technology Demonstration**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Environmental Security Technology Certification Program (ESTCP),4800 Mark Center Drive, Suite 17D08,Alexandria,VA,22350-3605**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**72**	

COST & PERFORMANCE REPORT

ESTCP Project: UX-0034

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (continued)

Page

FIGURES

Page

FIGURES (continued)

TABLES

Page

ACRONYMS AND ABBREVIATIONS

2-D	two-dimensional
AEC	Army Environmental Center
AERTA	Army Environmental Requirements and Technology Assessments
DAS	data analysis system
DGPS	Differential Global Positioning System
DoD	Department of Defense
EM	electromagnetic
EMI	electromagnetic induction
EMIS	electromagnetic induction spectroscopy
EMMS	Electromagnetic man-portable system
ERDC	Engineer Research and Development Center
ESTCP	Environmental Security Technology Certification Program
FAR	false alarm rate
GPS	global positioning system
GTL	Geophysical Technology Limited
JPG	Jefferson Proving Ground
MMS	man-portable magnetometer system
MTADS	miltisensor towed array detection system
NAVEODTECHDIV	Naval Explosive Ordnance Disposal Technology Division
NRL	Naval Research Laboratory
P(det)	probability of detection
P(disc)	probability of discrimination
Pfa	probability of false alarm
Pfp	Probability of false positive
QA	Quality Assurance
QC	Quality Control
RFP	request for proposal
ROC	receiver operator characteristic
RTK-DGPs	real-time kinetic differential global positioning system
SD	stop dig
UXO	unexploded ordnance

ACKNOWLEDGEMENTS

We wish to acknowledge the contributions made by the U.S. Army Engineer Research and Development Center (ERDC), (Dr. Ernesto Cespedes, "Jay" Bennett, Diane Cargile, and Thomas Berry) in the data analysis and in support of the field demonstrations. We appreciate the logistic support and contribution by the Pacific Division Naval Facilities Engineering Command (PACDIV), (Jim Putnam, Ronald Boyd, Carlton Finley, and Doug Dang) and field support from the Naval Explosive Ordnance Disposal Technology Division (NAVEODTECHDIV), (Scott Steward, Mike Hollander, Peter Pham, Steve Lange, Denice Forsht, and "Penny" Lanham).

Finally, we wish to express our grateful appreciation to the Environmental Security Technology Certification Program (ESTCP) Program Office for their commitment and continual support to the UXO detection and remediation technologies.

Technical material contained in this report has been approved for public release.

1.0 INTRODUCTION

The Environmental Security Technology Certification Program (ESTCP) funded the Naval Explosive Ordnance Disposal Technology Division (NAVEODTECHDIV) (lead agency), the U.S. Army Engineer Research and Development Center (ERDC), and the U.S. Army Environmental Center (AEC) to design and conduct controlled demonstrations of advanced unexploded ordnance (UXO) detection and discrimination technologies at the U.S. Army Jefferson Proving Ground (JPG) in Madison, Indiana, during FY 00 and at the Island of Kaho'olawe in Hawaii during FY 01. At JPG, these technology demonstrations were conducted at three 1-hectare areas located near the test site used during the JPG Phase IV demonstrations. At Kaho'olawe, the demonstrations were conducted at two prepared sites. The demonstrations were designed to evaluate the capabilities of state-of-the-art technologies to detect, discriminate, and identify buried UXO in areas containing high concentrations of natural (magnetic rocks/soils) and man-made (munitions fragments) clutter. This report documents the results of these demonstrations and provides data to aid the government in selecting effective and efficient systems for UXO detection and discrimination in difficult magnetic sites such as those encountered at Kaho'olawe Island.

1.1 BACKGROUND INFORMATION

The Department of Defense (DoD) is involved in UXO site remediation efforts where rapid transition of advanced technologies can potentially improve UXO detection efficiency, save substantial sums of money by reducing false alarms, and significantly expedite the transfer of lands for re-use. One of the most prominent of these efforts is the ongoing UXO cleanup of the Kaho'olawe bombing ranges. The major difficulty with this site is that the significant magnetic anomalies from geologic sources and near-surface metal fragments make traditional magnetometer-based surveys impractical. Active electromagnetic induction (EMI) sensors such as the Geonics EM-61 and the GTL TM-5 EMU are the primary sensors being used by the contractors at Kaho'olawe. Even though these EMI sensors have proven more effective at this site than passive magnetometers, their detection performance at Kaho'olawe sites has not been quantified, and they have been subject to very high false alarm rates (FAR). Parsons UXB, the prime UXO contractor at Kaho'olawe, reports that as of November 14, 2001, they have detected 61,261 subsurface anomalies and, after digging, found that only 2.7% are UXO, 27% are false positives from geologic sources, and 70.3% are the result of buried metal from both UXO and non-UXO-related materials. It should be noted that it is not possible to evaluate the detection performance, probability of detection (P(det)), from these findings, since the actual number of buried UXO (ground truth) is not known. ESTCP funded this project to address the critical need for more effective and efficient UXO technologies at sites such as Kaho'olawe.

The first phase of this ESTCP project was conducted at JPG July through November in 2000. This phase involved three advanced EMI sensing system demonstrators—NAEVA Geophysics, employing the Geonics EM-63 multichannel, time-domain EMI system; the Naval Research Laboratory (NRL), employing the single-channel, time-domain electromagnetic man-portable system (EMMS); and Geophex Ltd. employing the multifrequency, frequency-domain GEM-3 system. A commercial UXO surveying firm, EODT, was contracted to conduct standard mag-and-flag surveys of the JPG test areas to compare the EMI systems' performance with conventional techniques. Results of this first phase of demonstrations are documented in Reference 1 and indicate that the EMI systems perform considerably better (higher detection rates, fewer false alarms) than standard mag-and-flag surveys, especially in areas containing high levels of magnetic clutter from geologic sources. Since the first phase results provided strong indications that these technologies can significantly reduce false alarms resulting from high magnetic permeability in the soils and rocks, it was decided to evaluate the three systems under the more realistic and difficult geologic conditions found in Kaho'olawe.

In addition to the three advanced EMI systems demonstrated during the first phase, the ESTCP Program Office agreed to allow Geophysical Technology Limited (GTL) to demonstrate its advanced EMI sensor system, the TM-5 EMU, as part of the second phase tests at Kaho'olawe. GTL provided its own funding to participate in this phase, and ESTCP agreed to fund the additional costs associated with monitoring the GTL field surveys and evaluating their performance. It should be noted that GTL has participated in previous demonstrations conducted at JPG and was among the top performers in several test scenarios, including the small UXO sites (e.g., grenades and submunitions) (Reference 2). In addition, GTL has considerable operational experience with the TM-5 EMU in Kaho'olawe live sites.

For baseline comparisons with technologies currently used at Kaho'olawe sites, Parsons-UXB conducted standard EM-61 digital surveys as well as EM-and-Flag surveys using the EM-61 and the TM-5 EMU in a real-time detection/discrimination mode.

The focus of this demonstration project was to evaluate these advanced EMI technologies under realistic and difficult field conditions in order to quantify their detection, discrimination, cost, and production rates while operating at several areas within Kaho'olawe with varying degrees of target/clutter densities and magnetic noise levels. The purpose of this report is to aid managers of UXO cleanup projects as well as regulators and other stakeholders to make informed decisions concerning the capabilities, costs, and risks associated with applying these technologies to their site-specific UXO remediation problems.

1.2 OFFICIAL DOD REQUIREMENT STATEMENT

This project addresses the Tri-Service Environmental Quality Research, Development, Test, and Evaluation Strategic Plan, UXO requirements, and more specifically, the U.S. Army requirement A(1.6a) (Unexploded Ordnance (UXO) Screening, Detection, and Discrimination) and describes the FY 99 Army Environmental Requirements and Technology Assessments (AERTA). This Army requirement has been ranked as the highest priority user need in the Environmental Cleanup Pillar. In addition, this project addresses the UXO detection and discrimination requirements and recommendations described in the Defense Science Board Task Force Final Report on UXO Clearance and Remediation published in 1998 and will provide data to support the development of more accurate estimates of the overall DoD UXO environmental remediation costs.

The advanced technologies demonstrated as part of this effort address all aspects of the requirements for land-based, man-portable buried UXO detection and discrimination systems. The results of these demonstrations will be used to quantify the capability of state-of-the-art systems to detect, discriminate, locate, and identify buried targets. The performance of the advanced systems was compared with the baseline capability demonstrated by the on-site contractor, Parsons-UXB.

1.3 OBJECTIVES OF THE DEMONSTRATIONS

The primary technical objective of this demonstration project was to evaluate the detection and discrimination capabilities (including production rates and costs) of advanced UXO systems in difficult magnetic clutter environments such as those encountered at Kaho'olawe, Hawaii. At JPG, three test areas were prepared to present a limited range of conditions to the various demonstrators to identify scenarios where one technology may be better suited than the others. At Kaho'olawe, one 90 m by 111.1 m (1-hectare) area and 10 (not necessarily contiguous) 30 m by 30 m test grids within the Kaho'olawe Quality Assurance (QA) Range were prepared to present a limited range of target/clutter/topography/vegetation/magnetic background conditions to the various demonstrators.

The evaluation objectives for the demonstrations were as follows:

- To evaluate the demonstrators' detection and discrimination capabilities by surveying three 1-hectare areas within Jefferson Proving Ground, 30 m by 30 m grids, and one 1-hectare area within the Kaho'olawe QA Range under realistic target/geologic clutter/man-made clutter/topography scenarios and while operating as efficiently as possible (minimizing time, manpower, and costs).

- To evaluate the demonstrators' ability to analyze survey data in a timely manner and provide prioritized "dig lists" with associated confidence levels.

- To collect data on manpower and time required collecting field data necessary to produce their final products (prioritized dig sheets and georeferenced anomaly maps).

- To compare the performance of the advanced systems with the baseline mag-and-flag technologies at JPG and other technologies currently employed at Kaho'olawe.

- To provide high quality, well ground-truthed, georeferenced data for post-demonstration analysis and development of receiver operating characteristic (ROC) curves.

1.4 REGULATORY ISSUES

The principal regulatory issue affecting UXO detection and discrimination technologies is gaining confidence and approval from federal, state, and local regulators; stakeholders; and users. In addition, acceptance of these innovative technologies from agencies such as the U.S. Army Corps of Engineers and the Naval Facilities and Engineering Command is needed to ensure that future requests for proposals (RFP) for UXO cleanup projects will be written in a manner that will either sanction these technologies, or at least allow their inclusion in proposals for site work.

1.5 PREVIOUS TESTING OF THE TECHNOLOGY

Versions of the technologies demonstrated under this effort have been previously tested as part of other DoD and Army sponsored demonstrations, including the DARPA Clutter Experiment (FY 97), the Jefferson Proving Ground Phases II through IV Demonstrations, and a number of ESTCP-funded field demonstration projects. However, this ESTCP project represents the first set of controlled field experiments at an actual remediation site where these advanced technologies have been tested under realistic conditions that allowed for side-by-side comparison of detection/discrimination performance, production rates, and costs.

2.0 TECHNOLOGY DESCRIPTION

2.1 DESCRIPTION

The five electromagnetic induction sensing systems that participated in this ESTCP demonstration project consist of the following (in chronological order): the Geonics Ltd. EM-63, a multichannel time domain EMI sensor operated by personnel from NAEVA Geophysics; the GTL TM-5 EMU; a multiperiod time domain EMI sensor operated by GTL and Parsons-UXB Technology personnel; the Geophex Ltd. GEM-3, a multichannel frequency domain EMI sensor system operated by Geophex Ltd. Personnel; the NRL EMMS adjunct to the multisensor towed array detection system (MTADS) system, a single channel time domain EMI sensor operated by personnel from NRL with processing support from AETC Corp.; and the Geonics EM-61, a single channel time domain EMI system operated by Parsons-UXB. Each of the five sensors was integrated into a man-portable platform that included data acquisition/storage that merged the sensor data with position data collected by differential global positioning system (GPS) receivers.

At JPG, in addition to the EMI surveys conducted by three systems—GEM-3 (shown in Figure 1), EMMS (shown in Figure 2), and EM-63 (shown in Figure 3)—magnetic surveys of the three areas were conducted by NRL with a combination of the MTADS vehicular-towed magnetometer array and the man-portable magnetometer system (MMS) (shown in Figures 4 and 5), and by EODT Technology, Inc. (a commercial UXO services firm under contract to the U.S. Army Corps of Engineers, Engineering and Support Center, Huntsville), using the Schonstedt hand-held GA-52Cx magnetic gradiometer (shown in Figure 6). The MTADS/MMS platforms collected georeferenced total magnetic field data over the three test areas. The purpose of the MTADS/MMS survey was to collect a more complete data set to support post demonstration analysis and to identify/quantify any performance improvements resulting from adding magnetometer information to the EMI data. The Schonstedt GA-52 Cx is an analog magnetic gradiometer that provides only an audio signal to the operator when it senses a disturbance in the magnetic field (most likely caused by a buried ferrous object). The operator is then responsible for interpreting the strength and spatial extent of the audio signal to determine if it corresponds to an UXO-sized object; if so, he places a plastic pin flag at the estimated location of the object. EODT personnel were provided samples of emplaced ordnance and were instructed to disregard any buried object that they determined to be smaller than the smallest emplaced munitions (20 mm projectiles). ERDC personnel then surveyed each flagged location to produce the georeferenced mag-and-flag maps included in this report. The purpose of the Schonstedt survey was to establish a baseline for detection performance, cost, and production rate for comparison with the advanced EMI systems.

At Kaho'olawe, in addition to the five EMI systems (EM-63, TM-5 EMU, GEM-3, EMMS, and EM-61) Parsons Technology operated the TM-5 EMU and the EM-61 in a field discrimination (EM-and-Flag) mode. In this mode, the systems did not record digital sensor data, and the only permanent record consists of the identified UXO locations that were marked by a separate GPS survey crew. These sensor systems are shown in Figures 7 through 12 conducting surveys at the Kaho'olawe QA range.

Figure 1. GEM-3 Operated by Geophex Ltd.

Figure 2. EMMS Operated by NRL.

Figure 3. EM-63 Operated by NAEVA.

Figure 4. MMS Operated by NRL.

Figure 5. MTADS Operated by NRL.

Figure 6. GA-52Cx Operated by EODT.

Figure 7. EM-63 Operated by NAEVA.

Figure 8. TM-5 EMU Operated by GTL.

Figure 9. GEM-3 Operated by Geophex, Ltd.

Figure 10. EMMS Operated by NRL.

Figure 11. EM-61 Operated by Parsons.

Figure 12. TM-5 EMU Operated by Parsons.

2.2 STRENGTH, ADVANTAGES, AND WEAKNESS

The following paragraphs represent a summary of the perceived, claimed, and documented capabilities of each sensor employed by the technology demonstrators.

2.2.1 Geophex GEM-3

The advantage of the GEM-3 system is claimed to be its ability to rapidly collect multiple channels of complex frequency domain EMI data over a wide range of audio frequencies (30 Hz to more than 20 kHz), which allows it to perform what Geophex Ltd., the developer of the system, calls Electromagnetic Induction Spectroscopy (EMIS) on buried objects (Reference 5). EMIS potentially provides a method to discriminate UXO targets from natural and man-made clutter objects by means of their unique, complex (inphase and quadrature) frequency responses. The GEM-3 system required high-accuracy position information to perform the discrimination. For the search requirements of the current project, Geophex had to rely on GPS position information, which results in significant position errors and sparse data sets. Finally, it was observed that the GEM-3 system is still undergoing development and the sensor design, platform, data acquisition system, and analysis approaches have not been finalized or optimized.

2.2.2 Geonics EM-61

The Geonics EM-61 system is a single channel time-domain metal detector. It is the most widely applied EMI technology for UXO detection surveys (Reference 7). The system is available with different coil configurations and the one used at Kaho'olawe by Parsons-UXB incorporated the large 1 m by 1 m coils. As is the case with other time-domain systems, the transmitter coil generates a pulsed primary magnetic field in the earth, which induces eddy currents in nearby metallic objects. The eddy current decay produces a secondary magnetic field measured by the receiver coil. By taking the measurement at a relatively long time after the start of the decay, the designers of the EM-61 predict that the currents induced in the ground have fully dissipated and only the current in buried metal objects is still producing a secondary field. The responses are recorded and displayed by an integrated data logger.

2.2.3 NRL EMMS

The EMMS is derived from the MTADS development effort and thus incorporates many of its sensing, navigation, and data analysis system (DAS) advances demonstrated and documented in a number of ESTCP-funded field demonstrations. The specifications and performance improvements incorporated into the version of the EMMS demonstrated at Kaho'olawe are fully described in the ESTCP report, "Man-Portable Adjuncts for the MTADS" (Reference 6). Based on a modified version of the commercially available Geonics EM-61 (with the 0.5 m by 1.0 m transmitter coil), the most widely used EMI system for UXO detection applications, the EMMS sensor is expected to detect UXO to the maximum depths of the objects emplaced at Kaho'olawe. Coupled with the accuracy of the MTADS-derived, digital inclinometer/GPS system, the EMMS is expected to produce high quality georeferenced EMI data. A potential limitation of the EMMS is the single channel of data available, which may limit the discrimination performance compared to what can ultimately be achieved by multichannel systems.

2.2.4 Geonics EM-63

The advantage of the EM-63 lies in its ability to rapidly collect multiple channels of time-domain information at each survey point. The EM-63 collects up to 26 geometrically spaced time gates covering the time decay response in the range from 180 microseconds to 25 milliseconds after pulsing of the transmitter coil (Reference 3). Since the shape of the decay curve is dependent on the shape, size, orientation, and metal composition of the buried object, the EM-63 provides data that may be used to detect, discriminate, and identify the buried UXO targets, and to potentially reject responses from geologic materials and metallic clutter. The EM-63 is a commercially available sensor (produced by Geonics Ltd., which also manufactures the EM-61) and has been ruggedized for field use.

2.2.5 GTL TM-5 EMU

The GTL TM-5 EMU is a multiperiod, time-domain EMI system integrated with real-time processing that is claimed to provide automatic background leveling capability for enhanced detection and discrimination capabilities in sites containing high levels of magnetic interference (Reference 4). Unlike the other EMI systems tested during this demonstration, the TM-5 is a monocoil sensor with one element acting as both transmitter and receiver. The TM-5's transmitted waveform is referred to as "multiperiod" because it consists of a wavetrain with a single, longer pulse followed by three shorter pulses with the same length, all of which are repeated at a rate of approximately 1,200 Hz. The long pulse is four times wider than each of the short pulses. The decay period following each of the four pulses is sampled twice, with the specific details of gate timing, periods, and the method of combining them for analysis being proprietary information of MineLab Electronics and not available for publication. GTL has taken the sensor and electronics provided by MineLab and integrated advanced processing and positioning information to produce the TM-5 EMU specifically for UXO detection applications. The TM-5 EMU can perform UXO processing in real time, or the data can be recorded in digital form and post-processed to allow for more in-depth analysis and interpretation. Unfortunately, descriptions of the real-time and post processing techniques are also proprietary to GTL and not releasable. In spite of GTL's assurance that these details would be fully disclosed in return for being allowed to participate in the government's Kaho'olawe demonstrations, such information has not been provided. Thus, claims for the TM-5 EMU's automatic background leveling capabilities and its superior ability to operate in highly conductive and/or magnetic environments cannot be fully evaluated.

2.3 FACTORS INFLUENCING COST AND PERFORMANCE

Data on factors that influence the overall cost and performance of each of these systems in actual UXO remediation efforts were collected as part of this field demonstration effort and include the following:

- Equipment setup and calibration time and man-hour requirements
- Time and man-hour requirements to survey the demonstration test areas
- Downtime due to system malfunctions and maintenance requirements
- Re-acquisition/resurvey time and man-hour requirements

- Accuracy of georeferenced maps and prioritized dig lists with respect to:
  - Probability of detection (P[det])
  - FARs (probability of false positive [Pfp], FAR, Total FAR)
  - Discrimination capability (P[disc])
  - Identification capability
  - Target location accuracy.

The demonstration work plan (Reference 8) includes detailed descriptions of the methods and metrics used to evaluate each of the cost and performance factors.

3.0 DEMONSTRATIONS

3.1 DEMONSTRATION AT JEFFERSON PROVING GROUND

3.1.1 Background

The selection criteria for the three JPG demonstration areas are detailed in the Site Preparation Plan (Reference 3). The selection of the test areas was driven by the main demonstration objective, which was to evaluate the performance of advanced EMI technologies in the presence of magnetic noise from geologic sources and in different terrains. In addition, the three sites were seeded with varying concentrations of inert UXO and man-made clutter items.

3.1.2 Site/Facility Characteristics

The three 1-hectare areas within JPG were selected to provide the demonstrators with varying degrees of natural magnetic clutter and terrain difficulty. Area 1 was selected because it contains very high magnitude magnetic anomalies from geologic sources that cover a fairly large area, as shown in Figure 13. The long magnetic anomaly (grid square L-5) appearing near the center of Area 1 represents variations from the background mean of +150 nT to –100 nT, as measured by the MTADS system during previous JPG surveys. Area 1 has sparse tree/shrub coverage and its topography includes rolling terrain and ditches. It was seeded with the largest concentration of target and clutter items, a substantial number of which were placed within the high magnetic background locations. Area 2, also shown in Figure 13, was chosen because it has a significant number of magnetic geologic anomalies. In Area 2 the magnetic anomalies are more compact and lower in magnitude (+ 35 nT), thus providing a different clutter problem from that of Area 1. The topography in Area 2 also includes rolling terrain and a small ravine. Area 2 was seeded with a smaller number of target and clutter items than Area 1. Area 3 was chosen because it has very low amplitude magnetic anomalies from geologic sources and very flat terrain. This area has a variation from the mean background of only + 6 nT. Area 3 was seeded with the fewest UXO target and clutter items.

3.1.3 Performance Objectives

The scope of this demonstration was not intended to be a competition in which the government declares an overall winner. Its purpose was to collect sufficient information from this limited range of test scenarios to quantify the advantages and disadvantages of each of the three technologies so that they may be properly applied to specific UXO cleanup problems. The immediate goal of this effort was to collect the data needed to identify appropriate technologies to transition to the Kaho'olawe environments where natural (magnetic rocks/soils) and man-made (munitions fragments) clutter have rendered cleanup operations using conventional technologies both expensive and ineffective. A longer term objective of this demonstration is to provide high-quality, georeferenced data to support sensor development and improvements in UXO analysis technologies.

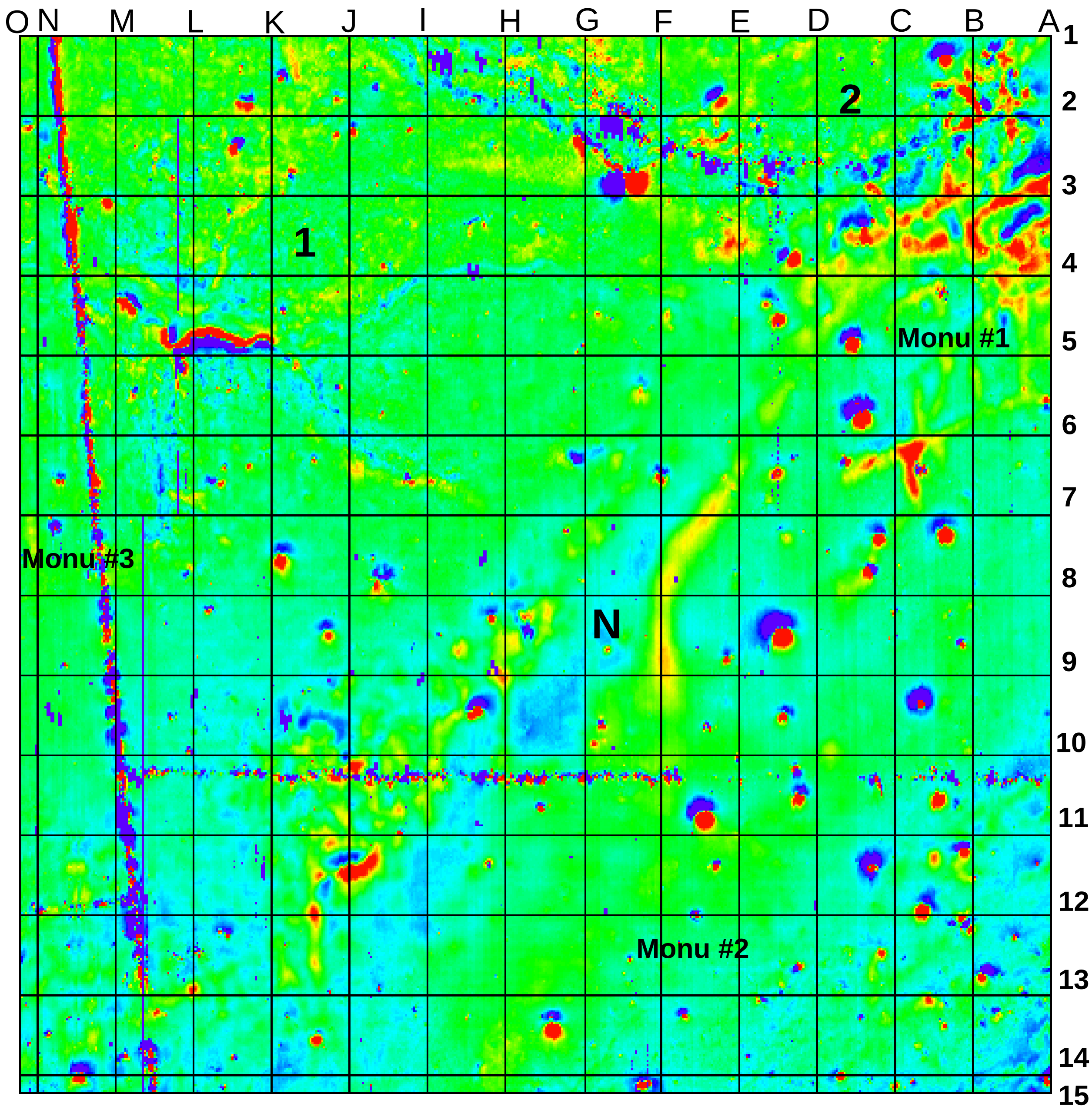

Figure 13. JPG Site Map Showing Magnetic Anomalies in Areas 1 and 2.
(Magnetic data collected by MTADS system [NRL] and provided by AETC)

The goals of this demonstration were to:

- Evaluate demonstrators operating in realistic target, geologic clutter, man-made clutter, and topography scenarios. Criteria for evaluation are:
  - Detection, discrimination, and identification capabilities based on prioritized dig lists produced from on-site data analysis.
  - Manpower, time, and costs required to produce on-site dig lists.
  - Additional detection and discrimination capabilities based on off-site, post-demonstration analyses.

- Provide baseline data for comparison of advanced EMI technologies with traditional mag-and-flag.

- Archive high-quality, ground-truthed, georeferenced data for broader use in the UXO technology development community.

3.1.4 Physical Setup and Operation

Descriptions of the inert UXO targets and the clutter items used for this demonstration are included in the Site Preparation Plan (Reference 3). Photographs, descriptions, dimensions, and emplacement information of each target and clutter item are available as part of the ground-truthed information in CD form from the ESTCP Program Office. Briefly, the UXO targets ranged from 20 mm projectiles buried near the surface to 155 mm projectiles buried up to 1.2 m below the surface. UXO items were degaussed prior to emplacement. Clutter items emplaced ranged from small (less than 0.5 kg) to large (up to 5 kg) munitions fragments and included large magnetic rocks and man-made clutter such as horseshoes and metal banding. A 2 m by 2 m area around each planned target location was surveyed with a G-858 magnetometer to detect and remove any metallic objects prior to emplacing an inert UXO target.

Samples of each of the UXO targets emplaced were made available to each demonstrator prior to arriving on site for signature collection and system training, and additional samples were available at the demonstration site for calibration purposes. Unlike the previous JPG tests, the clutter items were not made available to the demonstrators for signature collection and system training. A 2 m-long by 0.75 m-wide by 0.75 m-deep trench in the calibration area (Reference 1) located near Test Area 2 was made available to demonstrators for system calibration and checkout purposes.

Three resurveyed, first-order control points located within the original JPG 40-acre site were made available for demonstrators to set up GPS base stations. The primary reference monument is located near the southwest corner of Test Area 3 and was used as the reference point for all site preparation and demonstration activities. This marker was brought up to first-order accuracy during the site preparation activities, and updated coordinates were provided to the demonstrators prior to the scheduled demonstrations. Two other monuments were also resurveyed to first-order accuracy and made available to the demonstrators. One was designated Monument #1 (see Reference 1) and is located within Test Area 2 near its south boundary. The other is designated Monument #3 and is located approximately 40 m southwest of Test Area 1.

The four corners of each test area were surveyed by the government and marked with a metallic marker (rebar) driven flush with the ground for use by the demonstrators as fiducial markers to check/correct their position information. Plastic pin flags were placed at 5 m increments along the perimeter of each of the test areas to assist in maintaining proper lane spacing.

The demonstration test areas were mowed as part of the site preparation activities during June 2000. Prior to starting surveys, the first demonstrator (Geophex) inspected the test areas and determined that additional mowing was not required for their survey activities. The site was mowed for a second time prior to arrival of the second demonstrator (NRL). No additional mowing was conducted until the completion of the mag-and-flag surveys during November 2000.

3.1.5 Sampling Procedure

The Demonstration Work Plan describes the procedures required for each of the demonstrations. Demonstrators were responsible for developing their specific survey plans (lane spacing, sampling rate, number of channels recorded, calibration methods, etc.) and these procedures, together with their analysis techniques, are described in Reference 3.

Each of the demonstrators was allotted one 10-day period (Monday through Wednesday of the following week) from August 14, 2000, through September 20, 2000, to complete their surveys and submit the required on-site dig sheets. Each workday could extend to a maximum of 10 hr on site.

3.1.6 Analytical Procedures

The evaluation factors, metrics, products, and procedures related to this demonstration are described in the Demonstration Work Plan and are as follows:

1. Equipment setup, calibration time, and man-hour requirements

2. Actual survey time and man-hour requirements for each of the three test areas

3. Downtime because of system malfunctions and maintenance requirements

4. Reacquisition/resurvey time and man-hour requirements (if any)

5. Actual data processing/analysis time and man-hour requirements (all to be performed on site)

6. Prioritized dig lists with associated confidence levels

7. Discrimination capability (ability to separate detected anomalies into UXO and non-UXO objects)

8. Identification capability ability to classify UXO targets by class (e.g., mortar, projectile) and type (e.g., 152 mm)

9. Predicted target location accuracy (including depth estimates)

10. Georeferenced anomaly maps

11. Probabilities of detection

12. False alarm rates

To determine and document the first three items, government on-site representatives tracked and recorded the number of personnel and the time spent performing each task. Adequate rest and lunch/dinner breaks were provided, and these times were not included in the performance metrics calculations. If, during the analysis of the data, the demonstrator determined that he needed to resurvey any part of the test areas or any previously detected anomalies, all setup, calibration, survey, downtime, and reacquisition times and man-hour requirements were recorded individually (as in items 1 though 3), but were compiled separately as reacquisition/resurvey time (item 4).

To evaluate item 5, the government required that all data processing and analysis tasks required to produce items 6 through 10 be conducted in the JPG office trailer, and no data be taken off site until these items were submitted to the on-site government representative. Demonstrators were responsible for providing all computer hardware, software, and support equipment needed to produce the required analysis products.

Development and evaluation of (previously listed) items 6 through 10 are as follows:

- Each demonstrator was required to combine the electromagnetic (EM) sensor data with the GPS position information to develop two-dimensional (2-D) anomaly maps of each 1-hectare area. These maps, with the corresponding digital geophysical sensor data, were analyzed to identify all detected anomalies that could potentially be a buried UXO target for each of the three areas. These anomalies were then tabulated into one preliminary dig sheet for each test area. The objective of this phase was to include as many anomalies in these lists as required to ensure as high a P(det) as possible for the full range of UXO targets considered.

- Each anomaly in each list was further analyzed to develop the final prioritized dig sheets, as illustrated in Table 1. The demonstrators were asked to refine the location (x, y) and estimate the burial depth (z) of each object, to attempt to separate (discriminate) UXO from clutter items, to identify UXO by class and type (if possible), and to rank the list in the following descending order: UXO—high confidence, UXO—medium confidence, UXO—low confidence, Clutter—low confidence, Clutter—medium confidence, and Clutter—high confidence. The list was required to include predicted ordnance class and size (e.g., mortar/81 mm) for all anomalies declared as UXO with high and medium confidence levels, and, if possible, UXO orientation (azimuth and inclination).

- Each demonstrator was required to specify a threshold on each prioritized list and recommend that all objects at or above that threshold be excavated and those below be left in place. The goal of this step is to evaluate the demonstrators' ability to discriminate UXO targets from clutter. To add realism to this discrimination decision process, demonstrators were instructed beforehand that the following cost factors would be applied: 1) For every clutter item selected for "digging," a $200 cost penalty was assigned (the average cost of excavating items at actual UXO remediation sites). 2) To reflect the unacceptable risk of leaving UXO in the ground, a very high penalty was assigned if any detected anomaly that corresponded to a UXO target was erroneously declared as clutter and placed in the "no dig" portion of the list. As a result, if one or more UXO items were placed in the "no dig" portion of the list, it would be assumed that the grid (i.e., the entire 1-hectare area) has failed the quality assurance and/or regulatory acceptance, and a cost penalty equal to the cost of a resurvey would be assigned. Missed targets (anomalies too weak to be included in the lists developed in the description of equipment setup, calibration time, and man-hour requirements discussed previously) in each area were also assigned a cost factor equal to the cost of a resurvey, but it should be noted that they reflect a deficiency in the sensor rather than in the analysis and decision making process. Missed targets are also reflected in the less than 100% maximum P(det) achieved by each system and are documented in this report to aid regulators and managers in assessing residual risks associated with the various sensing technologies.

Table 1. Sample Dig List.

DIG LIST: 1 **Demonstrator:** EMMS **Test Area:** 1 **Including 20 mm ?:** NO

Ranking	Northing Meters	Easting Meters	Depth Meters	Type Ordnance/ Clutter	Confidence	Size/ Weight	Azimuth Degrees	Inclination Degrees	Class	Type
001	4309738.557	641594.2038	0.9144	Ordnance	High	Large	180	20	Projectile	152 Mm
.										
.										
.										
.										
.										
050	4309689.964	641519.4151	0.89042	Ordnance	Low	Small	-	-	Projectile	Unknown
.										
.										
165	4309700.031	641516.8877	0.82296	Clutter	High	Medium	-	-	Frag	-

Items 11 and 12 were calculated from the prioritized dig lists as follows: Maximum achievable P(det)s for each area were calculated as the number of items in the entire list that correspond to emplaced UXO targets (even though they may have been misclassified as clutter) divided by the actual number of UXO targets emplaced in that site. Note that in order to be declared a correct detection, the declared object location must be within a 1-m radius of the actual emplaced target location. The operating (single point) P(det) was determined by calculating the number of actual UXO targets included in the list at or above the threshold described in the previous paragraph. Similarly, the operating (single point) FAR was calculated as the number of clutter items above the dig threshold. A ROC-like curve was developed by the government by varying the dig threshold until the maximum P(det) was reached and computing P(det) and FAR at each increment. Performance comparisons between systems include using the ROC-like curves to determine FAR at the P(det) required for Kaho’olawe Tier II clearance (P(det) = 85-%) and also using the single point performance (P(det) and FAR) of the mag-and-flag surveys as a baseline.

After each demonstrator had submitted the dig sheets described above, the timing for the analysis tasks was stopped and he was to be given the opportunity to reanalyze the data to develop prioritized dig sheets that take into account only targets larger than 20 mm projectiles (20 mm projectiles were assumed to be clutter for this portion of the evaluation). These dig sheets were to be submitted to the government representative prior to leaving the JPG site. However, because of a late start and ensuing hardware problems that required additional delays for collecting additional calibration, the first on-site demonstrator (Geophex Ltd.) was unable to complete all the required data analysis in the allotted time. They requested, and ESTCP approved, a deviation from the Work Plan requirement that all processing be conducted on site. As a result, only the initial sets of dig sheets (EM only including all targets) were required to be submitted prior to departing the site. In addition, the last demonstrator (NAEVA) did not deploy their computer workstations to the JPG site and subsequently requested and received approval from the ESTCP office to perform the processing off site. As a result, the integrity of the on-site costs analysis was compromised and affected the overall cost evaluation included in Section 6 of Reference 10. NAEVA transmitted the field survey data off site for processing and was able to submitted a set of dig sheets prior to departing the site, so a comparison of the detection and discrimination performance is still viable.

After all on-site analysis products had been submitted, the demonstrators were provided with magnetometer data collected by MTADS and asked to reanalyze their data off site using this additional information to develop final prioritized dig sheets for each test area.

3.1.7 Performance Assessment

In accordance with the Demonstration Test Plan, each demonstrator was responsible for determining the best method for employing his system to: (1) ensure full coverage of each test area, (2) collect high-quality sensor data to support detection and discrimination requirements, (3) achieve high production rates, and (4) minimize man-hour requirements and costs. All demonstrators were able to complete the field surveys within the allotted time periods (see Reference 1). Figures 14 through 22 show the georeferenced anomaly maps produced by each of the systems used during these demonstrations.

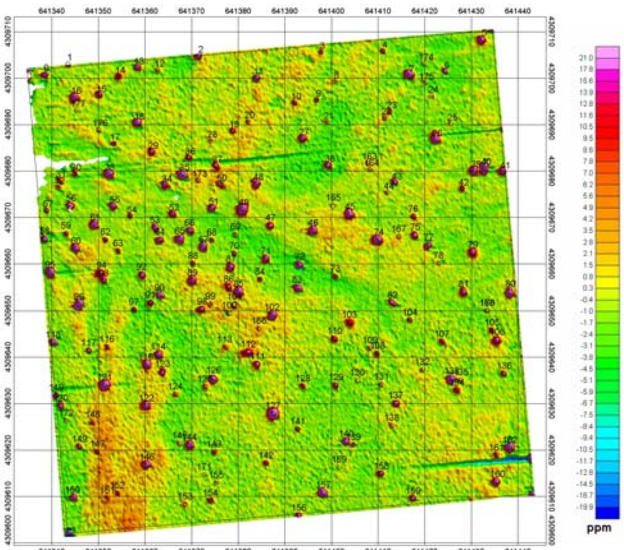

Figure 14. Geophex Ltd. GEM-3 Survey Results of Area 1.

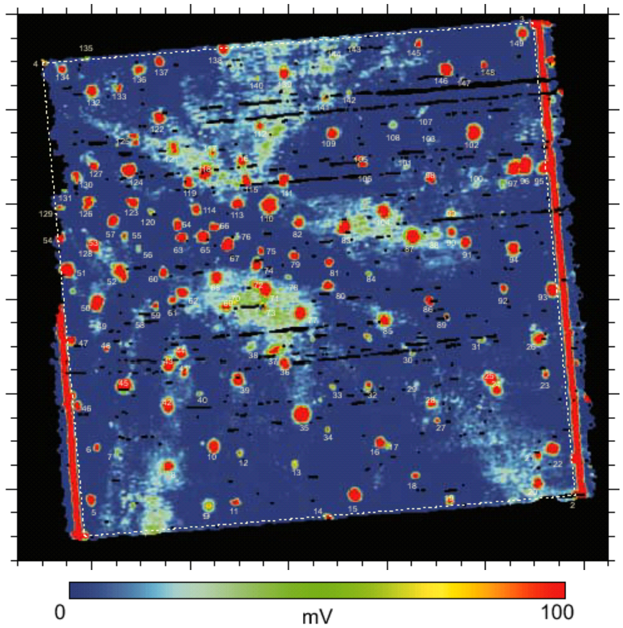

Figure 15. NRL EMMS Survey Results of Area 1.

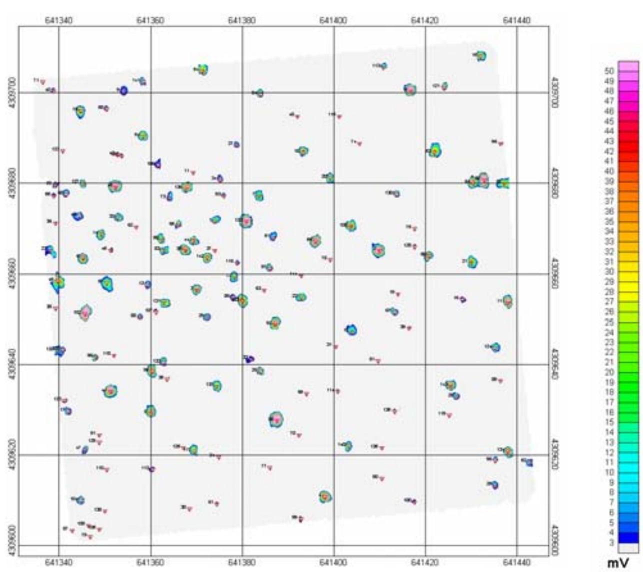

Figure 16. NAEVA EM-63 Survey Results of Area 1.

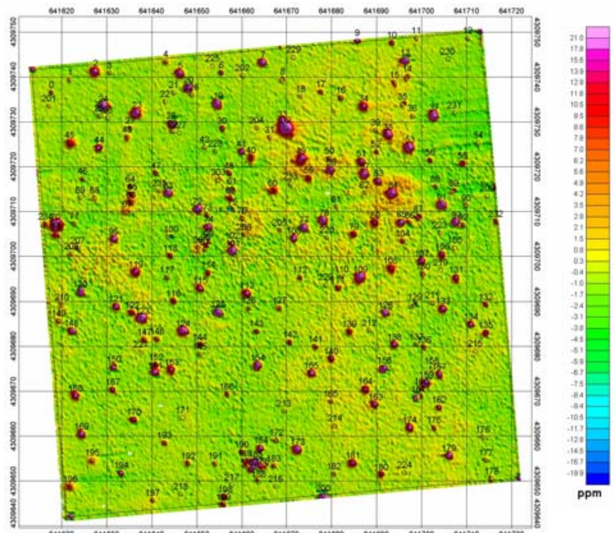

Figure 17. Geophex Ltd. GEM-3 Survey of Area 2.

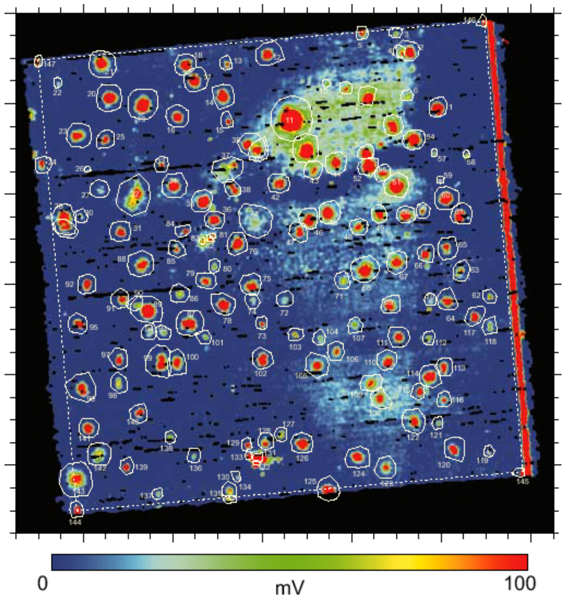

Figure 18. NRL EMMS Survey of Area 2.

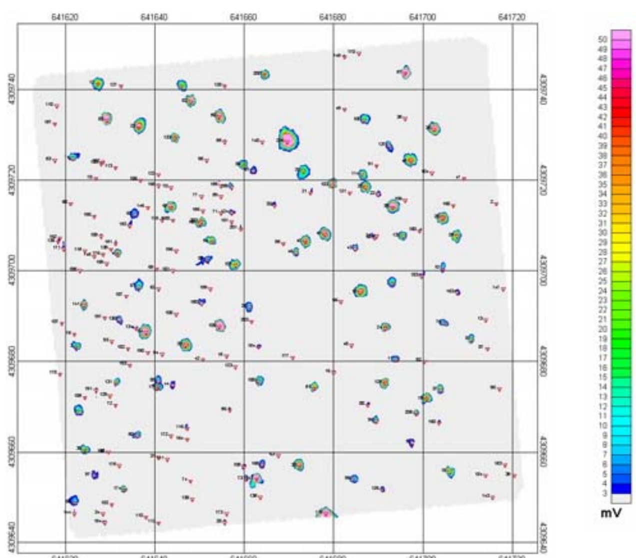

Figure 19. NAEVA EM-63 Survey of Area 2.

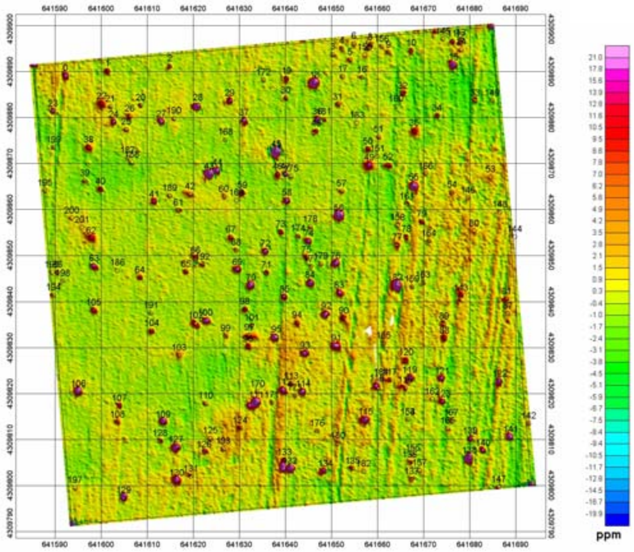

Figure 20. Geophex Ltd. GEM-3 Survey of Area 3.

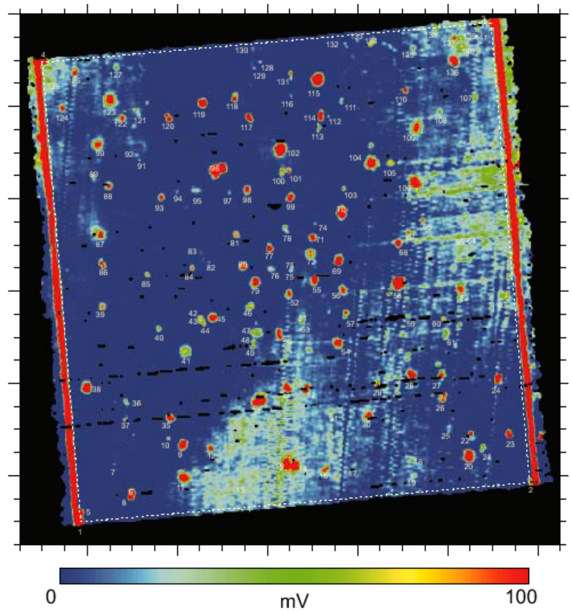

Figure 21. NRL EMMS Survey of Area 3.

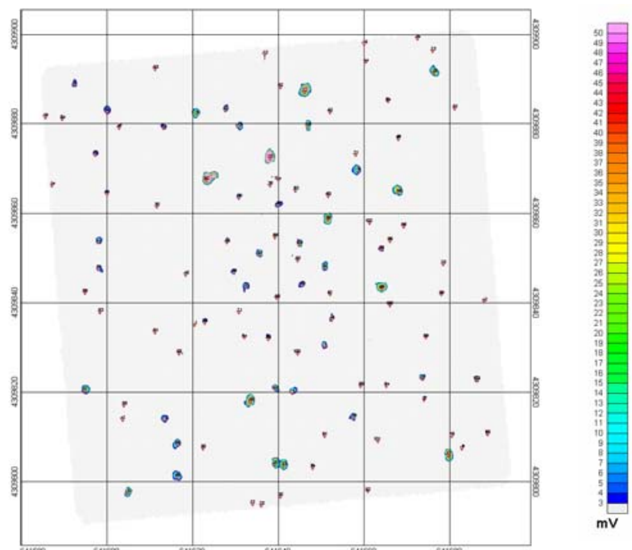

Figure 22. NAEVA EM-63 Survey of Area 3.

The results of the mag-and-flag surveys conducted by EODT (a commercial UXO firm under contract to the Corps of Engineers, Huntsville Engineering Center) were used as a baseline for documenting performance and cost improvements from the application of advanced EMI technologies. Government personnel using kinematic differential GPS (DGPS) equipment surveyed the locations flagged by EODT and the results are presented in Figures 23 and 24.

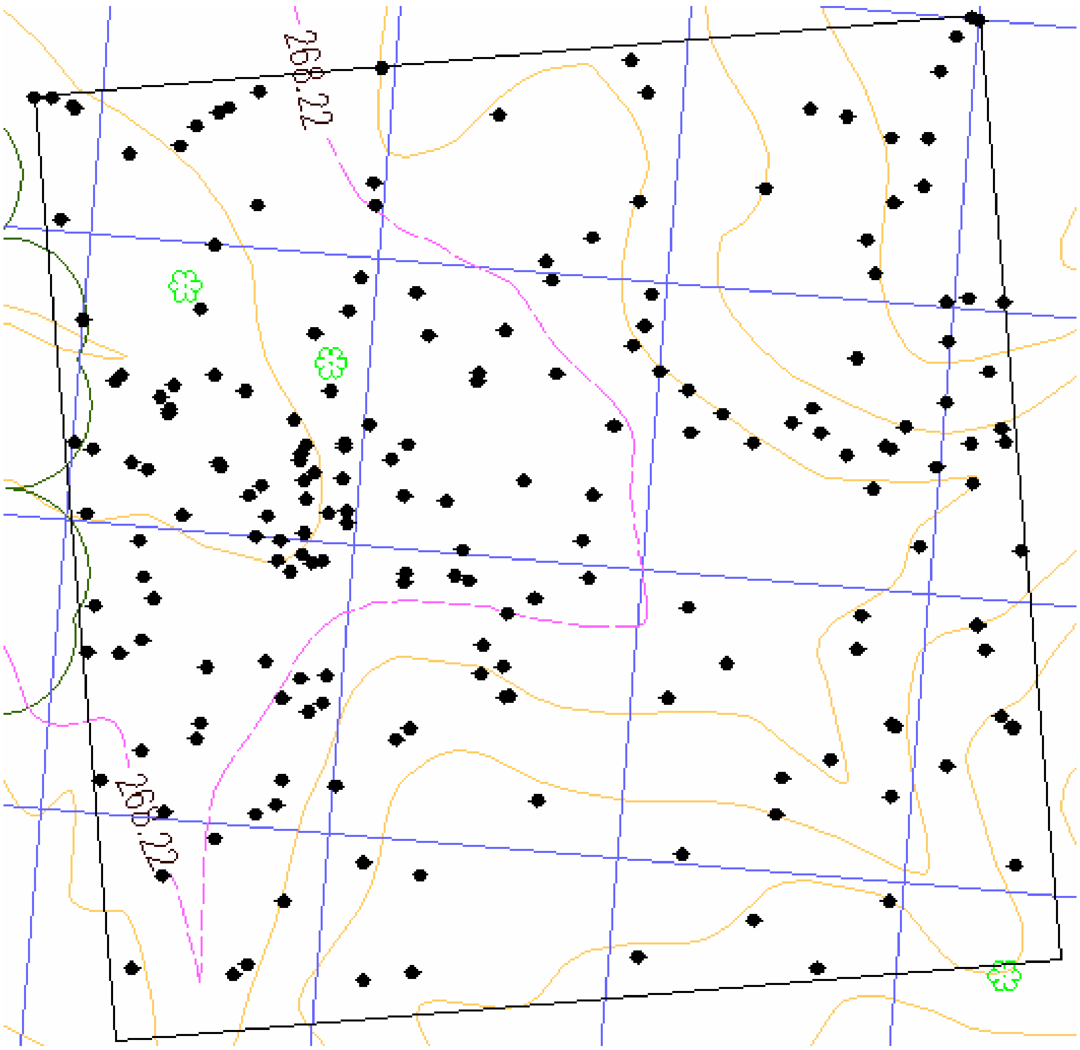

Figure 23. EODT Mag-and-Flag Survey of Area 1.

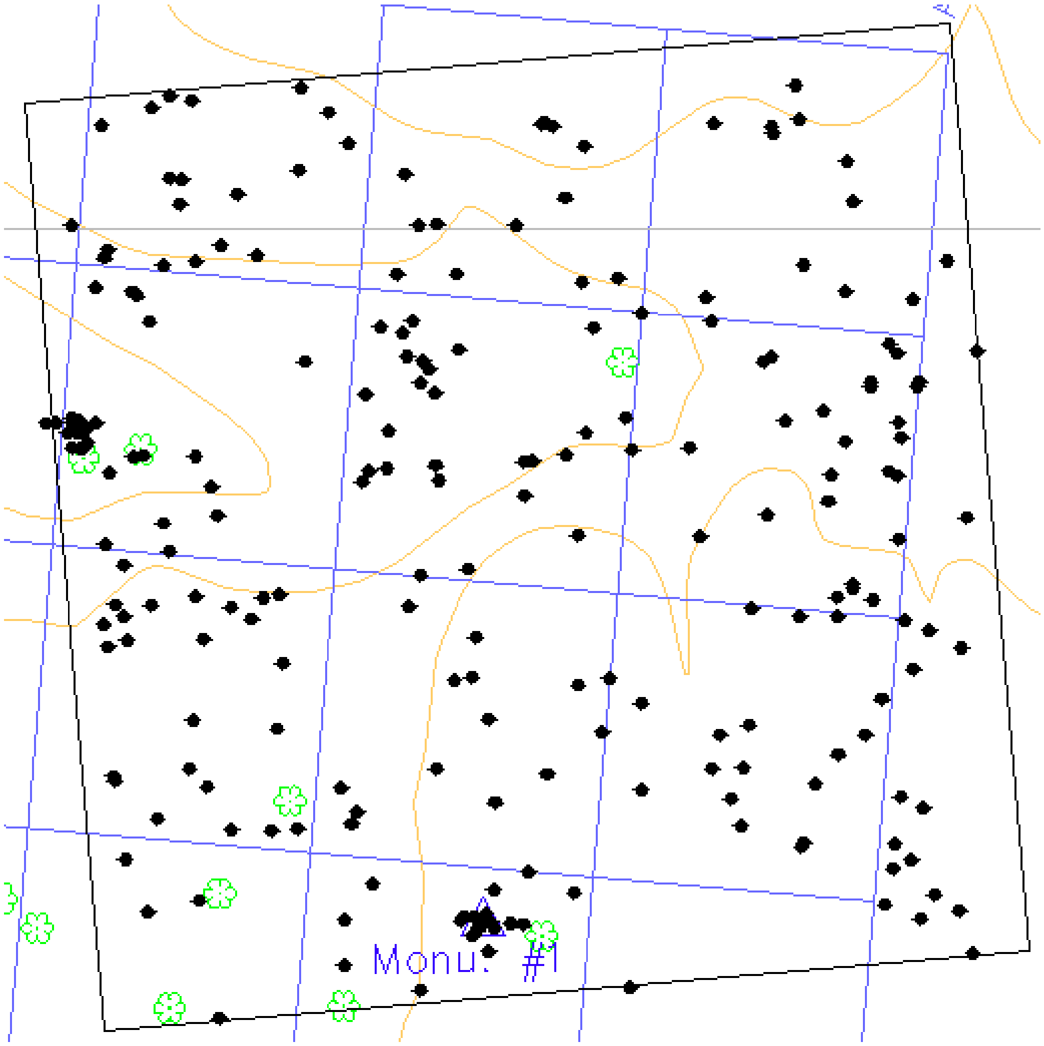

Figure 24. EODT "Mag-and-Flag Survey of Area 2.

3.1.7.1 Summary of Performance

One of the critical evaluation factors for this demonstration is the detection performance of the advanced systems. The metrics used to quantify the detection performance consist of the pseudo ROC curves, the single point P(det)/FAR, and the maximum achievable P(det). The methods used to estimate these metrics from the prioritized dig lists are described in detail in Section 4 of Reference 10. Briefly, the pseudo ROC curve, which graphically represents the target detection percentage versus the number of false alarms (or FAR in number of false alarms per hectare), is calculated by sequentially moving from the top of the prioritized dig list (i.e., the highest confidence UXO target declaration) and determining if each object on the list (whether classified as target or clutter) corresponds to an emplaced target location (a detection) or not (a false alarm). The single point P(det)/FAR performance is based on the point on the ROC curve that corresponds to the contractor-specified dig point on the prioritized dig list, and the maximum achievable P(det) is based on the highest point on the ROC curve. The placement of the "stop dig" (SD) point is shown as a triangle. These performance metrics are presented in the following graphs. The single point P(det)/FAR rate is shown as a colored triangle on the ROC curve, and the green diamond corresponds to the single point P(det)/FAR performance point of the mag-and-flag survey.

Figure 25 shows the detection performance of the three demonstrators based on the results of the on-site analysis that included all potential targets. The red traces show the performance results of the Geophex Ltd. GEM-3 system. The relatively flat slopes of these ROC curves indicate that the analysis performed on the GEM-3 data was not effective in discriminating UXO targets from clutter. The P(det) performance of the GEM-3 was superior to that of the standard mag-and-flag in the more difficult magnetic clutter environments of Areas 1 and 2 but did not demonstrate enhanced capability it in the low-noise environment of Area 3. In Areas 1 and 2, the single point GEM-3 P(det)/FAR performance failed to meet the 85% specified for the Kaho'olawe Tier II requirements. The GEM-3 was able to achieve 100% detection at all three sites but only at the expense of a significant number of false alarms.

In Figure 25, the blue traces show the performance results for the NRL EMMS. The steep early slope of the ROC curves indicates significant discrimination capability. The EMMS outperformed the mag-and-flag system at all three test areas, and the single-point performance points meet the Kaho'olawe Tier II requirements. Based on the maximum of the ROC curves, the EMMS did not achieve 100% detection at any of the three sites.

In Figure 25, the yellow traces show the corresponding performance results for the NAEVA EM-63 system. Again, the steep initial slopes of the ROC curves indicate significant discrimination capability. The P(det) performance of the NAEVA system was significantly better than mag-and-flag across all sites. NAEVA's ROC-based performance was very similar across Areas 1 and 3, and considerably lower for Area 2. The single-point performance points meet the Kaho'olawe requirements. The EM-63 system did not achieve 100% detection at any of the three sites.

The naturally occurring geologic magnetic noise and the emplaced magnetic rocks presented no problems to the three EMI systems. There were no false alarms attributable to anything other than metallic clutter in any of the submitted dig sheets, and analyses of the georeferenced maps

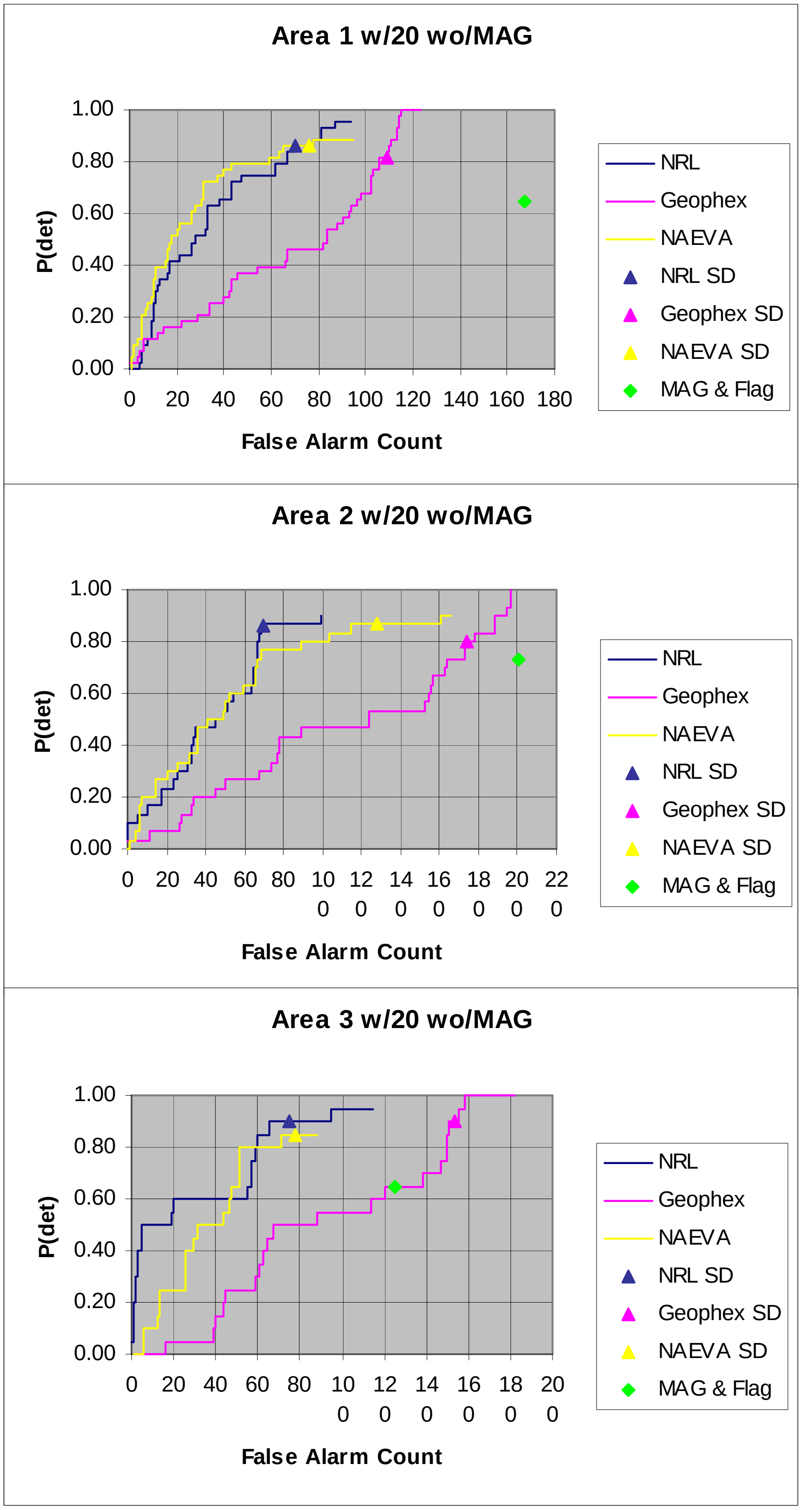

Figure 25. Detection Performance of GEM-3, EMMS, EM-63 System (On-Site Results).

show no discernible anomaly over any of the emplaced magnetic rocks. Overall, NRL and NAEVA demonstrated similar discrimination and FAR performance, and both were significantly higher than the demonstrated performance of Geophex Ltd. NRL's ROC performance, for Area 2 was slightly better than NAEVA's, while NAEVA's was very slightly better for Area 1. Overall, Geophex was the only system that demonstrated 100% P(det) at any of the three sites.

Table 2 provides point comparisons of the performance in Figure 25 (Reference 12).

Table 2. Assessment of P(det) Versus FAR with 20 mm Objects or Smaller without MAG.

	Area 1		Area 1		Area 2		Area 2		Area 3		Area 3	
Demonstrator	P(det)	FAR	SD P(det)	SD FAR	P(det)	FAR	SD P(det)	SD FAR	P(det)	FAR	SD P(det)	SD FAR
NRL	0.95	94	0.86	70	0.90	99	0.83	69	0.95	115	0.90	75
Geophex	1.00	123	0.81	109	1.00	196	0.80	174	1.00	182	0.90	154
NAEVA	0.88	95	0.86	76	0.90	166	0.87	128	0.85	88	0.85	78
M&F	**0.65**	**167**			**0.73**	**201**			**0.65**	**125**		

Figure 26 shows the detection performance of the three demonstrators based on the results of the off-site analyses that excluded objects that were estimated to be the size of 20 mm projectiles or smaller. The objective of this analysis was to determine the system performance based on the more commonly encountered, mid-sized (57 mm and larger) UXO targets. It should be noted that no comparisons with mag-and-flag results are included in these figures because analog magnetometers lack the capability to record the sensor data for re-analysis.

In Figure 26, the red traces indicate that the ROC-based performance of the GEM-3 system improved considerably from the on-site results shown previously in Figure 25. The off-site ROC curves have significantly steeper slopes (for all three areas) indicating much improved false alarm reduction capability. The operating P(det)/FAR points however, are much lower than in the previous set and, as a result, the GEM-3 operating P(det) was below 80% and failed to meet Tier II requirements for all three areas. The GEM-3 was able to achieve a max P(det) of 100% only on Area 1 and achieved only 81% max P(det) in Area 3. The significant decrease in operating and max P(det)s from the earlier results (where 100% max P(det) was achieved at all three areas) is difficult to explain. The objects dropped from the earlier dig lists as a result of this analysis consisted of a 105 mm projectile in Area 2 and a 60 mm mortar, an 81 mm mortar, and a 76 mm projectile in Area 3.

In Figure 26, the blue traces show the corresponding performance for the NRL EMMS system. Again, comparison of these results with those provided on site indicates significant improvement in ROC curve-based performance. In addition, the EMMS operating P(det)/FAR points improved substantially, particularly in Area 1 where 100% P(det) was obtained with only 55 false alarms. Maximum P(det) also increased slightly in the other two areas and exceeded the Tier II requirements.

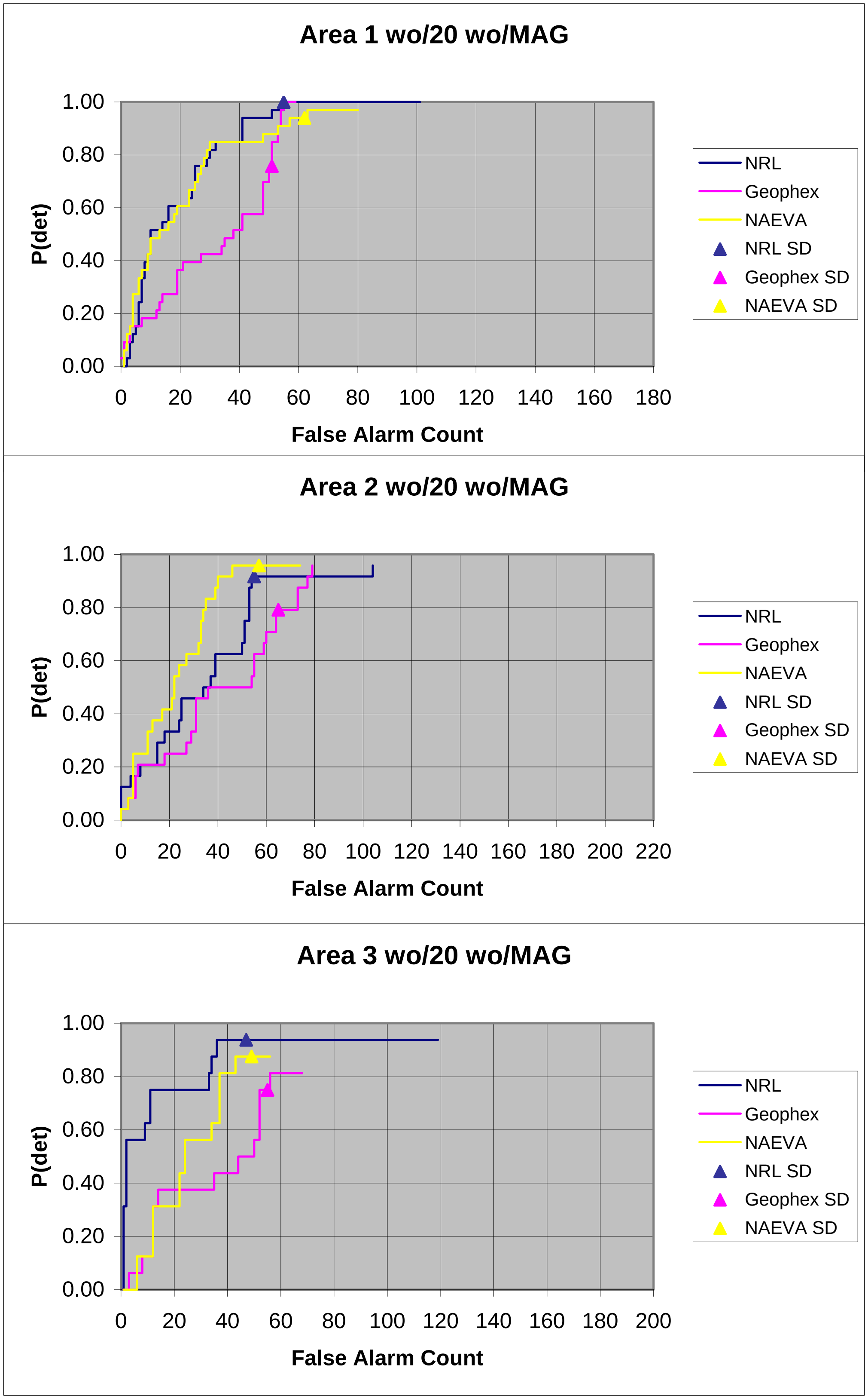

Figure 26. Detection Performance of GEM-3, EMMS, EM-63 System (Without 20 mm Targets).

In Figure 26, the yellow traces show the corresponding performance for the NAEVA EM-63 system. Again, a comparison with earlier results shows improvement in all factors. By far the greatest improvement is seen in Area 2 where the operating P(det) increased from 85% to 92%, while the corresponding false alarms were reduced from 128 to 56. Max P(det) increased significantly in Areas 1 and 2, and slightly in Area 3 and exceeded Tier II requirements, but again, NAEVA failed to reach the 100% maximum P(det) in all areas.

Overall, NRL and NAEVA demonstrated similar ROC-based performance, which was again significantly better than that demonstrated by the Geophex system. NRL demonstrated slightly higher performance in Areas 1 and 3, while NAEVA was slightly higher in Area 2. It can be concluded that excluding the small 20 mm targets resulted in significant performance improvements for the EMMS and EM-63 systems.

Figure 27 shows the performance of the three demonstrators when the MTADS mag data was added to the analysis and all targets were considered. The red trace shows that the overall GEM-3 performance improved very slightly from the EMI-only analysis presented in Figure 25. The mag-assisted ROC curve performance is slightly higher, but the FARs at the operating P(det)/FAR are still high. The operating P(det) did increase sufficiently to meet Tier II requirements for all three sites. The maximum P(det), however, was lowered by 10% in Area 2 and 5% in Area 3.

In Figure 27, the blue traces show that overall EMMS P(det) performance actually decreased with the addition of the mag data. Comparison of these results with those in Figure 25 show that, for all three areas, the ROC curve performance is lowered when the mag data is included in the analysis. In addition, the operating P(det)/FARs are significantly lower due to both a decrease in P(det) and an increase in the false alarms, and fail to meet Tier II requirements in all three areas. The maximum P(det) is also slightly lower in all three areas.

In Figure 27, the yellow traces show that overall EM-63 performance improved slightly over the results presented in Figure 25. The ROC curve performance improved slightly for Areas 1 and 3, and significantly for Area 2. False alarms for Area 2 were nearly reduced by a factor of 2 but only slightly reduced for the other areas. The operating P(det), however, is lower than those obtained without the mag data and fail to meet Tier II requirements for all three areas. The maximum obtainable P(det) was lower in all cases.

Comparing results of the three demonstrators indicates that any enhancements due to the addition of mag data are generally minor, and, in many cases, the addition of mag data actually degraded system performance. NAEVA demonstrated the largest performance improvement (Area 2) over the EMI-only analysis, and has the overall best ROC curve performance. NRL was slightly worse than NAEVA in Areas 1 and 2, and significantly worse in Area 3. The GEM-3 ROC curve performance improved slightly but is still well below that of the other two demonstrators. The largest impact observed was on the operating P(det) point, which improved for the GEM-3 while decreasing for the other two demonstrators. As a result, the GEM-3 was the only system meeting Tier II requirements. There appears to be no trend or reasonable explanation for the widely varying effects resulting from the addition of mag data in the analysis. For example, it

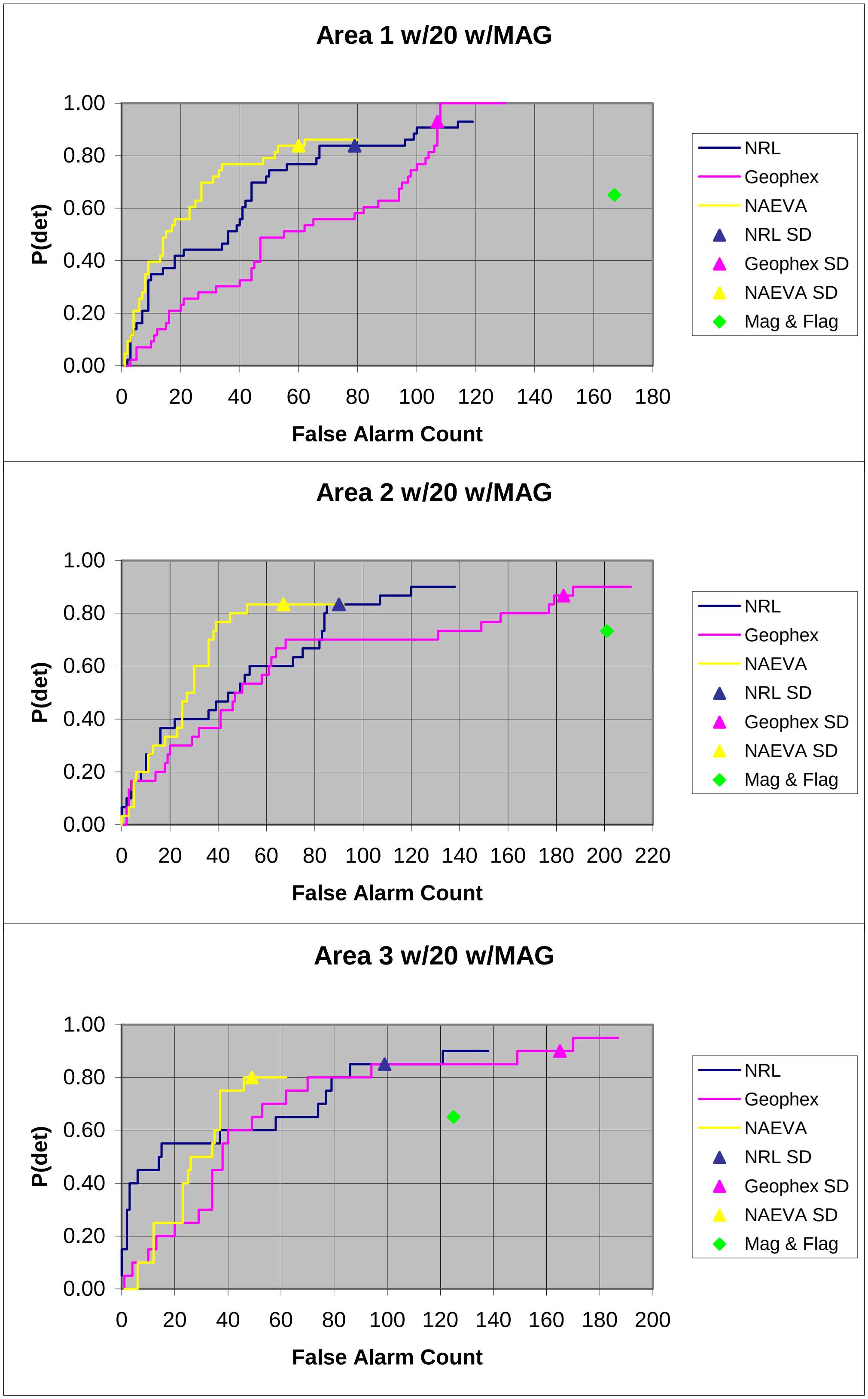

Figure 27. Detection Performance of GEM-3, EMMS, EM-63 System (MAG with 20 mm Results).

would be expected that mag data would significantly enhance the EMI results, especially in Area 3 where the geologic magnetic noise is minimal, but the data do not support this hypothesis. Table 3 provides point comparisons of the performance in Figure 27 (Reference 12).

Table 3. Assessment of P(det) Versus FAR with 20 mm Objects or Smaller with MAG.

	Area 1	Area 1	Area 1	Area 1	Area 2	Area 2	Area 2	Area 2	Area 3	Area 3	Area 3	Area 3
Demonstrator	**P(det)**	**FAR**	**SD P(det)**	**SD FAR**	**P(det)**	**FAR**	**SD P(det)**	**SD FAR**	**P(det)**	**FAR**	**SD P(det)**	**SD FAR**
NRL	0.93	119	0.84	79	0.90	138	0.83	90	0.90	138	0.85	99
Geophex	1.00	130	0.93	107	0.90	211	0.87	183	0.95	187	0.90	165
NAEVA	0.86	80	0.84	60	0.83	92	0.83	67	0.80	62	0.80	49
M&F	**0.65**	**167**			**0.73**	**201**			**0.65**	**125**		

Figure 28 shows the detection performance of the three systems against 57 mm and larger targets after demonstrators were allowed to integrate magnetometer data into their analysis of the EMI data. The purpose of this analysis is to quantify performance improvements in mid-sized UXO detection from magnetometer data under varying clutter conditions.

In Figure 28, the red traces show that the overall GEM-3 results with the mag data included is radically different from the EMI-only analysis presented in Figure 26. There seems to be very little correlation between the mag-enhanced and the EMI-only ROC curves, and these results are more closely correlated with the original on-site analysis that included the 20 mm projectile targets. It appears that the significant change seen in Figure 28 was due to a decision (or threshold) that resulted in the elimination of a large number of anomalies from the dig lists, and that this decision was reversed in the course of the subsequent mag-EMI analysis. Comparison of the ROC curves in these two figures shows that the mag-EMI analysis includes almost twice as many objects as the EMI-only analysis. The initial slopes of the two sets of curves are very similar, but the mag-EMI set levels off and continues to much higher false alarm counts. The operating P(det)s in Figure 28, while significantly higher, occur at much higher false alarm counts. The mag-EMI operating P(det)s exceed the Tier II requirements for all three areas, whereas all of the P(det)s from the EMI-only data failed to meet them. The maximum P(det)s are also considerably higher than for the EMI-only case and reach 100% for all three areas.

In Figure 28, the blue traces show that the overall EMMS performance improved very slightly from the EMI-only analysis presented in Figure 26. The ROC curve performance for all three areas is only slightly better than the EMI-only performance. The operating P(det) performance is worse for the mag-EMI case since, for all three areas, the operating P(det)s are slightly lower and occur at higher false alarm counts. Though lower, operating P(det)s still meet/exceed Tier II requirements in all three areas. The maximum achievable P(det) is slightly lower for Areas 1 and 3 and unchanged in Area 2. Overall, EMMS performance was not improved by incorporating the mag data.

In Figure 28, the yellow traces show that the overall EM-63 performance is almost identical to the EMI-only analysis presented in Figure 26. Comparison of the two sets of results reveals that

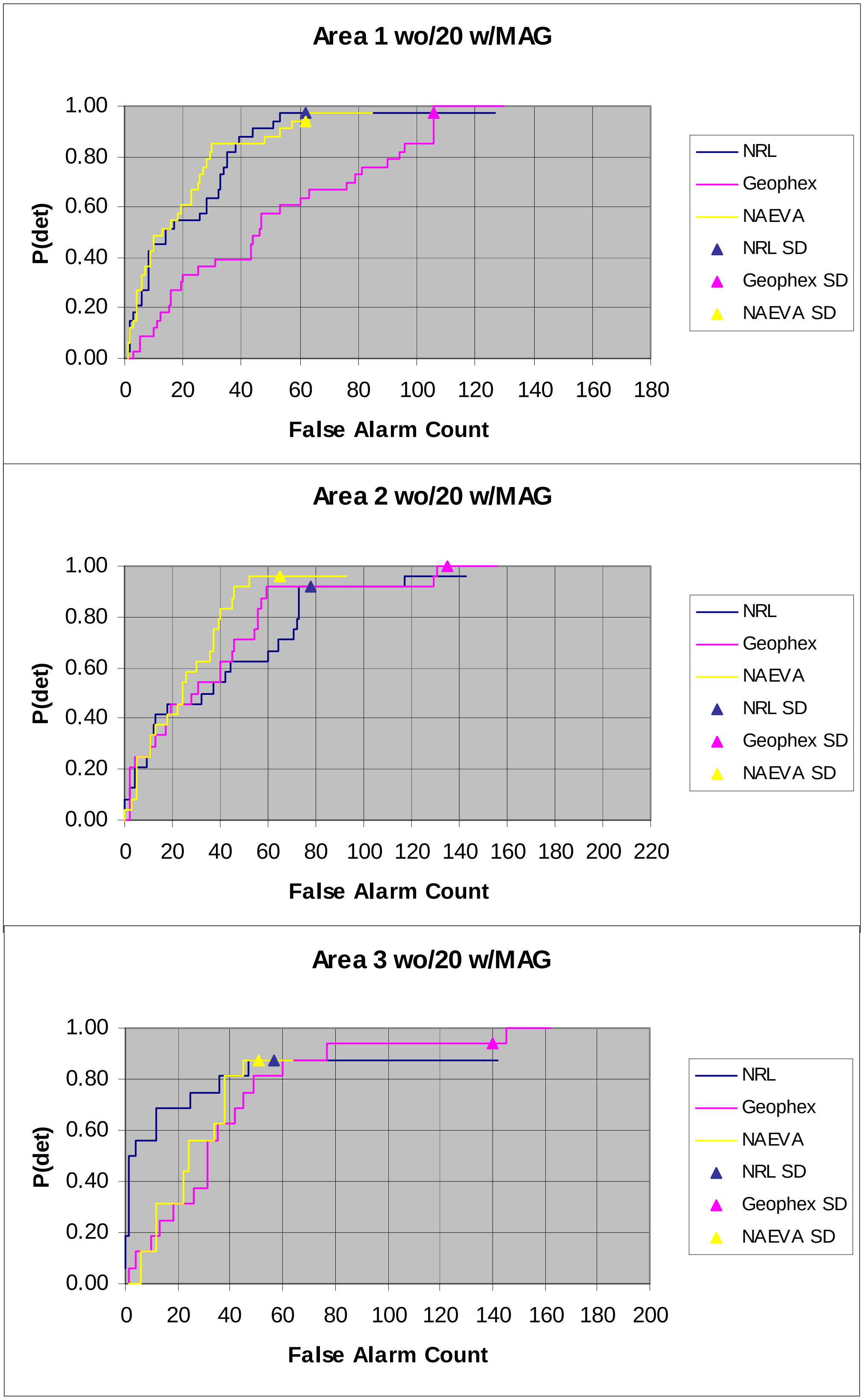

Figure 28. Detection Performance of GEM-3, EMMS, EM-63 System (MAG Without 20 mm Results).

the only change is the addition of a few false alarms to the high end of each of the mag-EMI ROC curves. As a result, the ROC curve performance, operating P(det)s, and maximum achievable P(det)s remain unchanged. Overall, the EM-63 performance was not significantly improved by the addition of the mag data.

3.1.7.2 Discrimination and Identification Results

The discrimination and identification capabilities of UXO systems greatly affect the cost and residual risks associated with any UXO cleanup operation. The assessment of these capabilities for the three advanced systems demonstrated at JPG can be found in Section 5 of Reference 10.

A comparison of the discrimination performance of the three systems shows that NAEVA demonstrated the best capability to reliably discriminate ordnance from clutter. NRL demonstrated considerably lower discrimination capability, and Geophex Ltd. demonstrated very poor discrimination capability. None of the systems could be considered to have demonstrated capability to identify ordnance items either by type or by class.

3.1.7.3 Depth Estimation Results

The ability of the demonstrators to estimate depth of the UXO targets is summarized in Table 4. These results indicated that, while the performance of each demonstrator varies significantly between test areas, the mean depth estimation errors were well within the desired 0.5 m allowable error. Overall, the NAEVA EM-63 systems achieved the best depth estimation accuracy, followed by the NRL EMMS, with the Geophex GEM-3 demonstrating the largest maximum and mean depths estimation error.

Table 4. Demonstrators' UXO Target Depth Estimation Performance.

Area	Demonstrator	Minimum Error (m)	Maximum Error (m)	Mean Error (m)	Standard Deviation
1	NAEVA EM-63	0.00	0.63	0.19	0.14
	NRL EMMS	0.01	0.65	0.23	0.17
	GEOPHEX GEM-3	0.01	0.81	0.20	0.18
2	NAEVA EM-63	0.03	0.72	0.24	0.19
	NRL EMMS	0.00	0.86	0.27	0.24
	GEOPHEX GEM-3	0.04	0.93	0.30	0.23
3	NAEVA EM-63	0.01	0.76	0.16	0.21
	NRL EMMS	0.01	0.37	0.16	0.10
	GEOPHEX GEM-3	0.02	1.10	0.31	0.27

3.1.8 Technology Comparison

Overall, discrimination, classification, and identification performance of all three systems was lower than expected and significantly lower than those demonstrated at JPG Phase IV. Some obvious reasons for the decreased performance are that, unlike at JPG IV, the demonstrators did not have prior access to the clutter items, and the systems were operated in a wide area search mode (rather than a point survey mode as JPG IV), which reduced the number of data samples (data density), the available signal strength levels, and the position accuracy. The detection performance, while acceptable, was poorer than demonstrated during previous demonstrations and suffered from high alarm rates.

The various advanced systems demonstrated varying degrees of maturity. The NRL EMMS, which demonstrated the highest degree of maturity and preparation, conducted the field surveys and on-site analysis with no problems. In contrast, the GEM-3 and EM-63 systems demonstrated a lower level of readiness during the field evaluations. The Geophex team experienced a sensor failure shortly after starting the survey and was forced to use a spare GEM-3, which used a different coil size. This required the collection of an entirely new signature library and on-site modifications to the analysis software. The NAEVA team arrived on site without the capability to perform the analysis tasks and was unable to demonstrate on-site processing performance.

3.1.9 Cost Assessment

The cost factors and penalties for false alarms and UXO targets left in the ground were applied to arrive at the cost comparisons summarized in Table 5. It should be noted that the data analysis costs are shown where available, but are not included in the totals since we were not able to determine the analysis cost for NAEVA.

Table 5. Cost to Survey.

	Cost to Survey
NRL	$9,654.10
GEOPHEX	$10,497.35
NAEVA	$10,816.20

3.1.10 Technology Implementation

The next step for these technologies is to transition to demonstration at prepared test sites on Kaho'olawe Island in the Hawaiian island chain. Kaho'olawe Island is rich in basalt, which results in a high magnetic background, making the detection of unexploded ordnance difficult. It is expected that the best performing technology will be rapidly transitioned to active cleanup operations at live sites throughout the island.

3.2 DEMONSTRATION AT KAHO’OLAWE ISLAND, HAWAII

3.2.1 Site/Facility Characteristics

Kaho’olawe Island consists of the summit of a single volcanic dome that reaches a peak elevation of 1,477 ft above mean sea level (Lua Makika point at the northeastern part of the island). It is one of the oldest of the main group of Hawaiian Islands and is separated from Maui by the 6.9-mile wide Alalake’ike Channel and from Lana’i by the 17.5-mile Kealaikahiki Channel. It is 11 miles long and 6 miles wide with an area of 28,776 acres. Kaho’olawe Island was used as weapons range and military training area from 1941 until 1990. Title X of the FY 94 Department of Defense Appropriations Act was enacted in November 1993 and directed the cleanup of ranges in Kaho’olawe Island. Title X allocated $400 million for UXO remediation starting in 1993 and required that Kaho’olawe be transferred to a Native Hawaiian sovereign entity not later than November 2003.

The island’s geology consists predominantly of basalt, hardpan, and sand. The magnetite-containing basaltic rocks and soils have precluded the use of magnetometers for UXO detection and have been the source of a significant number of false alarms encountered by the EMI sensors currently used by the UXO remediation contractors. Kaho’olawe’s surface features consist primarily of dry land vegetation and hardpan (Figures 29 and 30). The island’s climate is windy and very dry, averaging only 10 to 20 inches of rainfall per year (mostly on the eastern side of the island). The island has had a history of overgrazing by sheep and cattle, destruction of vegetation by goats, deforestation by settlers, and damage caused by target bombing and shelling by the U.S. military.

Figure 29. Overview of Kaho’olawe Showing Erosion of Terrain.

Figure 30. Overview of Kaho’olawe Vegetation and Hardpan Terrain.

3.2.2 Calibration and Demonstration Site Characteristics

The previously established Kaho’olawe UXO Remediation project’s quality control (QC) and QA ranges, shown in Figure 31, were selected for this project’s calibration and demonstration areas respectively. The calibration and demonstration sites were surveyed using electronic theodolites (Leica Model TCA 1102) and real-time kinematic differential global positioning system (RTK-DGPS) survey equipment. Both sites were oriented to true North and each corner of each range was staked with a ferrous rod and its coordinates recorded. The magnetic variation at the Kaho’olawe site is 9˚ 59’ East. Monuments near the calibration and demonstration areas were brought up to first order accuracy during the initial site preparation activities, and updated coordinates were provided to the demonstrators prior to the scheduled demonstrations.

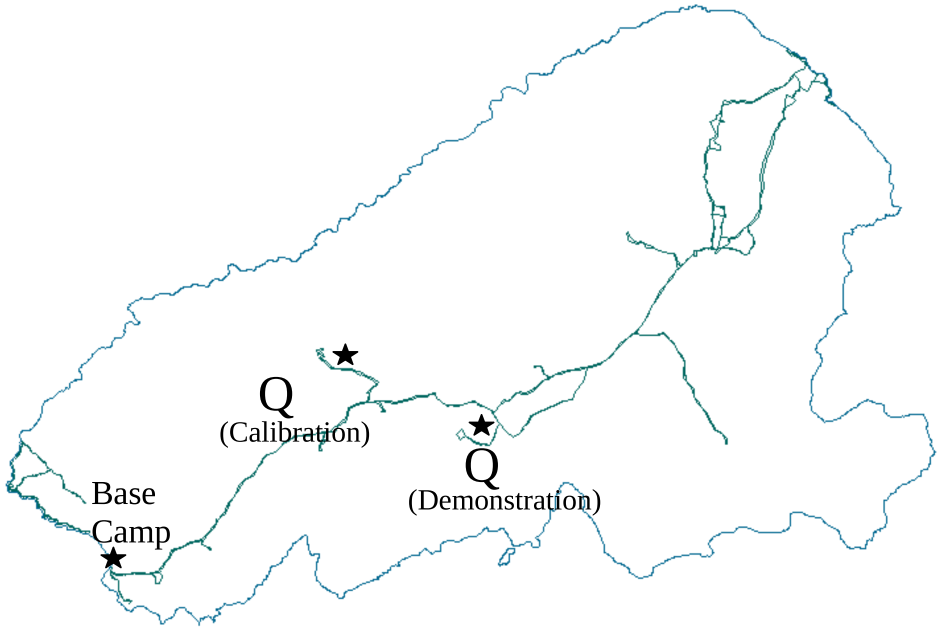

Figure 31. Kaho'olawe Map with Outline of Base Camp, Calibration, and Demonstration Ranges.

The calibration area consisted of three 30 m by 30 m grids (Figure 32) and was established to allow the demonstrators to conduct system calibration, signature collection, and algorithm development prior to participating in the blind tests conducted in the demonstration areas. The complete ground truth of all items emplaced in the calibration area was made available to each demonstrator prior to arrival on Kaho'olawe.

Figure 32. Calibration Area Within the QC Range.

The demonstration area was subdivided into three areas: Area A consisted of four 30 m by 30 m grids; Area B consisted of nine 30 m by 30 m grids and 3 partial grids (totaling 1 hectare); and Area C consisted of six 30 m by 30 m grids (Figure 33). Figure 34 shows the QA range looking from grid 5B to grid 1E.

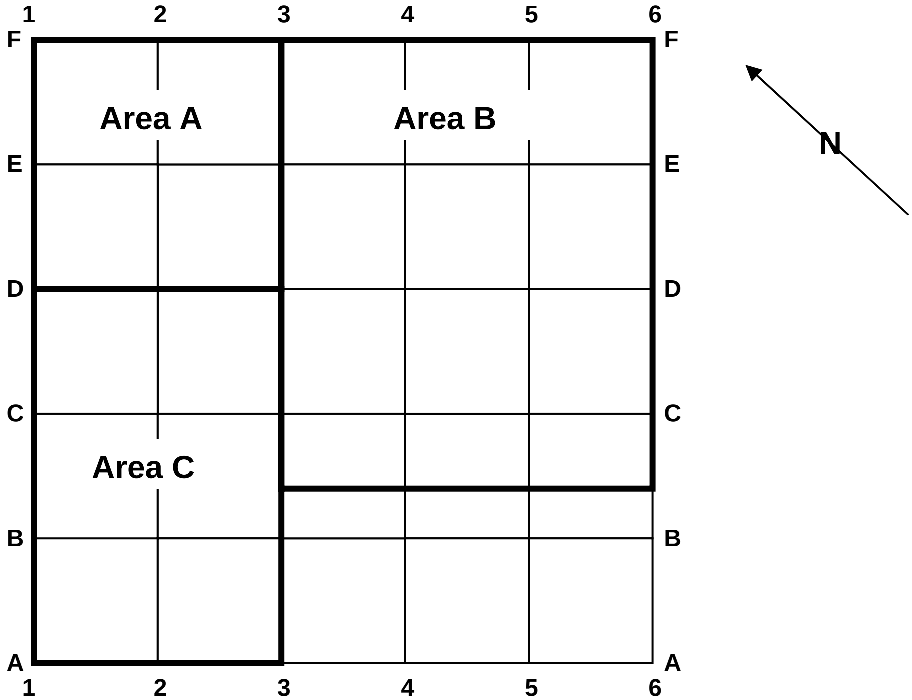

Figure 33. Layout of Demonstration Areas.

Figure 34. Photo of Demonstration Area Within the QA Range.

The target list consisted of the following 18 items. (The items in bold were taken from the Standard Test Sites Program stockpile, and the rest are from Kaho'olawe UXO cleanup operations.)

- **20-mm projectile**
- **40-mm projectile**
- **60-mm mortar**
- **81-mm mortar**
- 2.25-inch rocket
- **2.75-inch rocket warhead**
- 3-inch projectile
- 5-inch projectile
- **105-mm projectile**
- SMAW rocket
- LAAW
- BDU 3 practice bomb
- BDU 33 practice bomb
- MK 82 practice bomb
- MK-3 practice bomb
- MK-81 practice bomb
- MK-106 practice bomb
- MK-83 practice bomb

All unfired, inert UXO items were thoroughly degaussed either at NAVEODTECHDIV at Aberdeen Proving Ground, Maryland, or in the field prior to emplacement in order to remove any remnant magnetic moment. Each degaussed item was checked for remnant magnetization using a G-858 total field magnetometer. Target and clutter items were weighed, measured, and photographed, and the excavation crew emplaced the preselected target/clutter item at the location, orientation, and approximate burial depth specified in the Site Preparation Plan.

Prior to emplacing any targets, all UXO target emplacement locations (larger than 20 mm) were surveyed with a Geonics EM-61 HH and a Geometrics G-858 in order to record background noise prior to emplacement and also to allow removal of any metal objects in the vicinity (within a 2 m by 2 m area) of the target location. In addition, one 30 m by 30 m grid within the demonstration area that did not contain any pre-emplaced items was surveyed with both the EM-61 and the G-858. These pre-emplacement surveys are included in the data archive described in Appendix B of Reference 11. Figure 35 and Figure 36 show overviews of the clutter and target items in the calibration and demonstration areas, respectively. Detailed information on these items is included in the Site Preparation Plan.

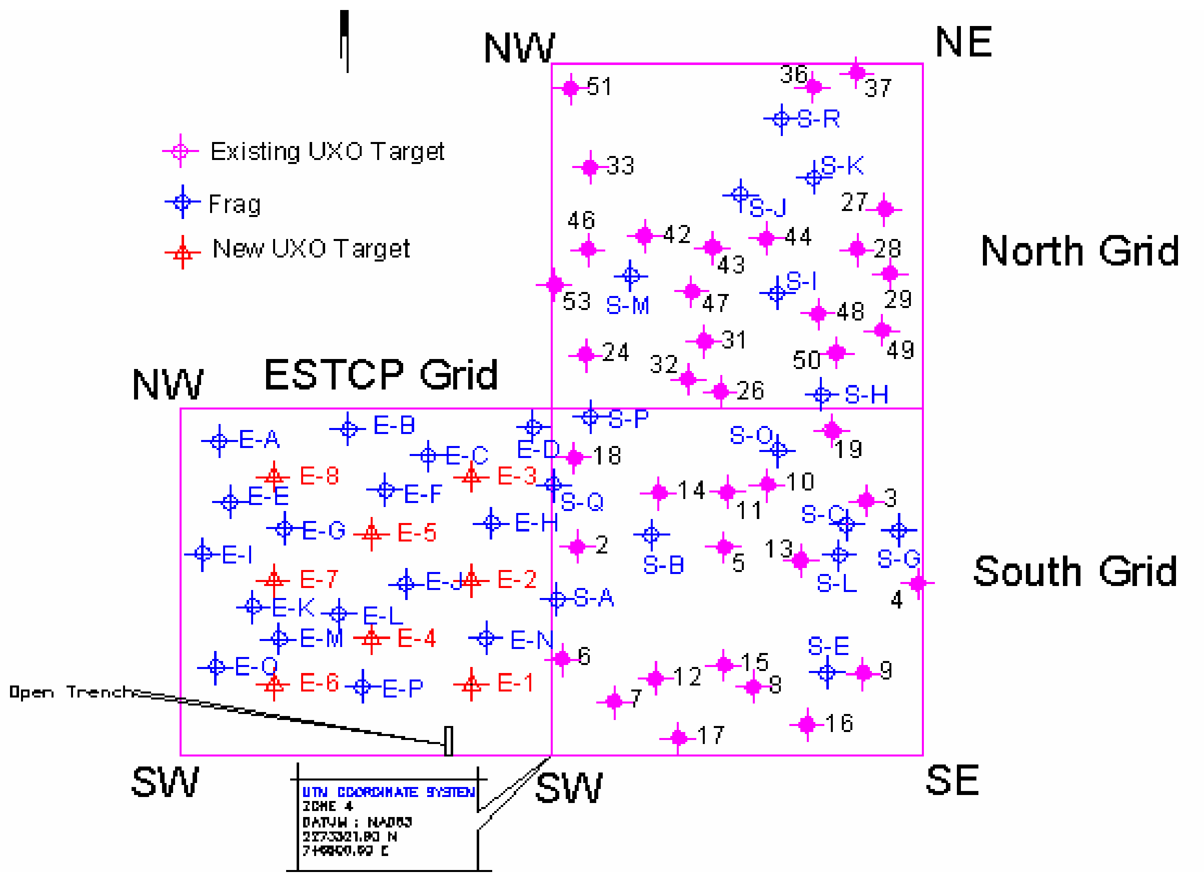

Figure 35. Kaho'olawe Calibration Area Target Emplacement Plan.

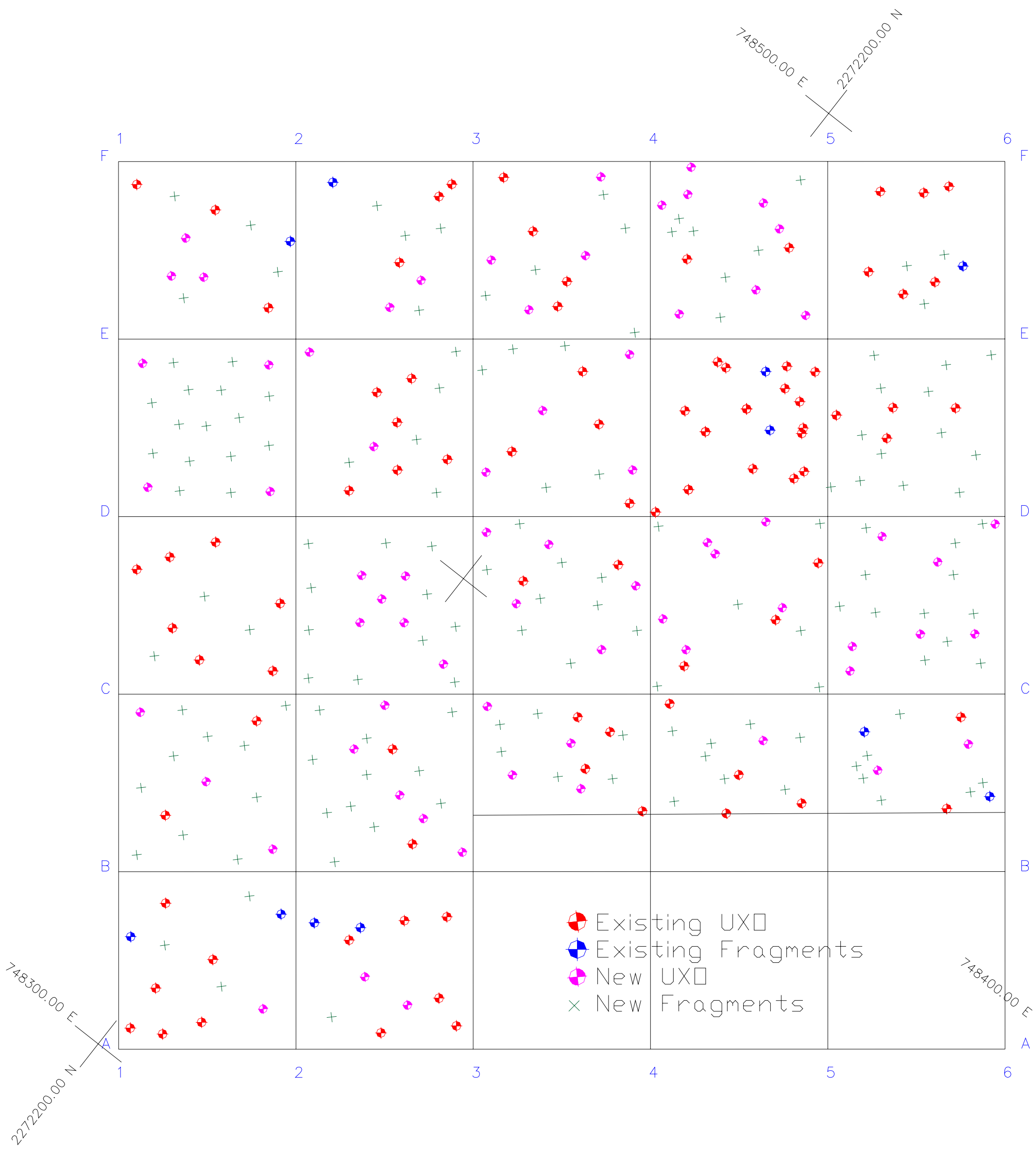

Figure 36. Kaho’olawe Site Map Showing Emplacement Items in Areas A, B, and C of the Demonstration Site.

3.2.3 Performance Objectives

The objective of this demonstration was to evaluate the performance of advanced EMI technologies in very difficult UXO target, clutter, and geologic noise scenarios such as those commonly encountered in Kaho’olawe UXO remediation efforts. As was the case at JPG during the first phase of this project, this demonstration attempted to evaluate the detection, discrimination, and identification capabilities of advanced UXO systems while simultaneously evaluating their production rates, manpower requirements, and costs.

The scope of this effort was to collect sufficient information from a limited range of test scenarios in order to quantify the advantages and disadvantages of each of the four EMI technologies so that they may be properly applied to specific UXO cleanup problems. The immediate goal of this effort is to quantify any performance and cost improvements resulting from the application of advanced EMI technologies so that they can be rapidly transitioned to Kaho'olawe operations where natural (magnetic rocks/soils) and man-made (munitions fragments) clutter have rendered conventional UXO cleanup efforts both expensive and ineffective. A longer-term objective of this demonstration is to collect and archive high-quality, georeferenced data to support future sensor development and improvements in UXO analysis technologies.

The performance objectives for this demonstration are as follows:

- To conduct surveys of one 1-hectare area and ten 30 m by 30 m areas within Kaho'olawe Island under very difficult but realistic target/geologic clutter/man-made clutter scenarios while operating as efficiently as possible (minimizing time, manpower, and costs).

- To analyze survey data in a timely manner (before leaving Hawaii) and provide "dig lists" that include detection, discrimination, and identification estimates with associated confidence levels as well as georeferenced anomaly maps. Note that unlike the first phase demonstrations at JPG, this demonstration phase did not require that the processing be performed on site.

- To detect and locate all buried ordnance while minimizing false alarms due to geologic and man-made clutter sources. Ordnance location accuracy was originally specified to be within a 0.5 m halo (horizontally) around the emplaced munitions. (This corresponds to excavating a 0.5 m horizontal radius hole centered at the declared location and striking any part of the UXO). Due to the highly cluttered environment, difficult site conditions (wind, terrain, heat) and uncertainties regarding the actual position of the pre-existing buried UXO targets, demonstrators' performance were also evaluated using a 1.0 m radius. (See Appendix C of Reference 11). Refer to Appendix D of Reference 11 for analysis of mean location error scatter plots for setting the halo at 0.5 m.

- To provide high quality, georeferenced data for post-demonstration analysis and development of ROC curves and for broader use in the UXO technology development community.

- To prepare technical reports to evaluate and document performance and to aid the government in selecting effective and efficient systems for UXO detection and discrimination in difficult magnetic clutter sites such as Kaho'olawe, Hawaii.

3.2.4 Physical Setup and Operation

Because Kaho'olawe Island contains numerous live UXO, rough terrain, and some areas of thick vegetation, the sites available for testing and demonstration are very limited. As a result, this project was restricted to operating in previously cleared sites that fell within the existing QC and

QA areas. Operating under these constraints, personnel from NAVEODTECHDIV and ERDC, with support from on-site Navy and contractor personnel, set up a calibration site at the QC area and a limited set of demonstration scenarios at the QA area.

The UXO targets for the calibration and demonstration areas consisted of the following: (a) inert UXO items previously used at Kaho'olawe for QA/QC purposes and whose age, weathering, condition can vary and (b) pristine inert munitions from the Standard UXO Test Site Program fabricated/procured to be as nearly identical as feasible. Several Demonstration grids were prepared so that the only UXO targets selected were from this set of standard test targets. Representative samples of ordnance emplaced at the calibration site were made available to demonstrators for viewing and for conducting free-air and buried measurements during the demonstration phase.

The QC grids were expanded to allow installation of the new UXO items obtained from the Standardized UXO Test Site stockpile. The demonstration areas were prepared by trimming and removing unwanted trees and tall grasses and by watering around any newly excavated areas to reduce evidence of site disturbance. No other physical alterations were made to the existing ranges as part of the site preparation activities conducted during September 3-26, 2001.

Descriptions of the inert UXO targets and the clutter items used for this demonstration are included in the Site Preparation Plan (Reference 9) which is available online as described in Appendix B of Reference 11. Photographs, descriptions, dimensions, and emplacement information of each target and clutter item are available as part of the information archived at this site. Briefly, the UXO targets ranged from 20 mm projectiles buried near the surface to 500 lb practice bombs buried to 2.5 m below the surface. Clutter items emplaced ranged from small (less than 0.5 kg) to large (over 5 kg) munitions fragments retrieved during UXO cleanup operations on Kaho'olawe.

Based on concerns raised during previous UXO detection and discrimination demonstrations, no unrealistic (fabricated) clutter items were used; instead, actual munitions fragments from past Kaho'olawe cleanup operations were used. Based on previous JPG experience, all inert UXO targets that had not been previously fired or air-dropped were demagnetized prior to emplacement to simulate the magnetic properties of the ordnance employed. Precautions were taken during target emplacement to minimize surface disturbances (e.g., "bathtub" effect) that could alert the demonstrators to the presence of a buried object.

3.2.5 Sampling Procedures

The Demonstration Work Plan describes the procedures required for each of the demonstrations. Demonstrators were responsible for developing their specific survey plans (including lane spacing, sampling rate, number of channels recorded, calibration methods, etc.) and these procedures, together with their analysis techniques, are described in Appendix E of Reference 11.

Each of the demonstrators was allotted one 3-week period (Monday through Thursday) during the period of September 17 - November 30, 2001 to complete the following activities:

- System calibration and signature data collection activities at the calibration area during the first week.

- Algorithm development and testing on Maui during the second week.

- Field surveys of the demonstration areas and data analysis/preparation of dig lists and georeferenced maps during the third week.

3.2.6 Analytical Procedures

The evaluation factors, metrics, products, and procedures related to this demonstration are described in the Demonstration Work Plan (Reference 8) and include the following information:

a. Equipment setup and calibration time and man-hour requirements

b. Actual survey time and man-hour requirements for each of the demonstration test areas

c. Downtime due to system malfunctions and maintenance requirements

d. Re-acquisition/resurvey time and man-hour requirements (if any)

e. Prioritized dig lists with associated signal strength and confidence levels

f. Discrimination capability (ability to separate detected anomalies into UXO and non-UXO objects)

g. Identification capability (ability to classify UXO targets by class [e.g., mortar, projectile] and type (e.g., 105 mm projectile)

h. Predicted target location accuracy (including depth estimates)

i. Georeferenced anomaly maps

j. P(det)

k. P(disc)

l. Pfp due only to the emplaced clutter items

m. FARs due to nonemplaced items (e.g., geology and/or unknown items)

n. Total FAR due to emplaced clutter items and nonemplaced clutter items

o. Operational costs.

The method for determining and documenting items a, b, and c consisted of the government on-site representatives tracking and recording the number of personnel and time spent performing each of these tasks. Adequate rest and lunch/snack breaks were provided and included in the performance metrics calculations. If the demonstrator determined that he needed to resurvey any part of the test areas or any previously detected anomalies, all setup, calibration, survey, downtime, and re-acquisition times and man-hour requirements were recorded individually (as in items a through c) but were compiled separately as re-acquisition/resurvey time (item d).

Development and evaluation of items e through i were as follows:

(1) Each demonstrator was to combine the EM sensor data with the GPS and other position information to develop 2-D anomaly maps (item i) of a test area consisting of four contiguous 30m by 30 m grids (Area A), a 1- hectare contiguous test area (Area B), and a test area consisting of six 30 m by 30 m grids (Area C). These anomaly maps, together with the corresponding digital geophysical sensor data were to be analyzed to identify all detected anomalies that could potentially be a buried UXO target for each of the test areas. All of these anomalies were to be tabulated into one preliminary dig sheet for each test area and were to include a suitable "signal strength" value determined and defined by each of the demonstrators (e.g., 1,100 ppm @ 930 Hz/quadrature phase for the GEM-3 system). Each demonstrator was required to submit a detailed description of and rationale for selecting this parameter as part of the prioritized dig list submissions described later in this document. The objective of the detection stage was to include as many anomalies in these lists as required to ensure as high a P(det) as possible for the full range of UXO targets considered.

(2) Each anomaly in each list was then to be further analyzed to develop the final dig sheets. The demonstrators were to refine the location (x, y) and estimate the depth (z) of each object, attempt to separate (discriminate) UXO from clutter items, identify UXO by class and type (if possible), and rank the list in the following descending order: UXO—high confidence; UXO—medium confidence; UXO—low confidence; clutter—low confidence; clutter—medium confidence; and clutter—high confidence. In addition, the list was to include predicted ordnance class and type (e.g., mortar/81mm) for all anomalies declared as UXO with high and medium confidence levels, and, if possible, UXO orientation (azimuth and inclination).

(3) Each demonstrator was then required to select a point (threshold) on each prioritized list where he would recommend that all objects at or above that point be excavated and the remainder left in place. We refer to this as the "stop-dig-point". The goal of this step was to maximize the number of UXO targets above the threshold while minimizing the number of clutter items. In order to add realism to this demonstration, the following cost penalties were to be applied to this product: For every clutter item selected for "digging," a $200 cost penalty was assigned (the average cost of excavating items at actual UXO

remediation sites). If the demonstrator included any UXO items in the "no dig" portion of the list, it was assumed that the area (i.e., either the 1-hectare area, four 30 m grids or six 30 m grids) would fail the quality assurance and/or regulatory acceptance and a cost penalty equal to the cost of a resurvey was assigned. Missed UXO targets, that is those not included in the dig lists either as clutter or UXO, were also penalized the cost of a resurvey. Note that the same resurvey penalty was applied only once for each area whether only one or any number of UXO items were left in the ground through mis-discrimination or by failing to detect it or a combination of both.

(4) In order to better approximate real-world UXO cleanup operations, the government offered to furnish ground-truth information of Area A within 24 hours after a demonstrator submitted prioritized dig lists for that specified portion of the test area. This procedure was intended to correspond to the additional information that is normally available to the UXO survey contractors when surveyed grids are excavated. It should be noted that ground truth corresponding only to anomalies included in the prioritized dig lists was provided. No information on missed targets was made available until the full ground truth was released after completing of the demonstrations. To provide this information as early as possible during the blind demonstrations, the procedure was as follows: At the beginning of the demonstration phase, each demonstrator was directed to survey Area A, consisting of four 30 m by 30 m grids. The demonstrator was then to proceed to develop a prioritized dig list of this area while his survey crew continued to collect field survey data of the remaining test areas. As soon as the demonstrator submitted the prioritized dig list, the government representative evaluated the results and provided ground truth information on the declared target and clutter items. The demonstrator could then use that information to modify his analysis and/or survey techniques during the remainder of the blind demonstration. The goal was for each demonstrator to have this information prior to starting the analysis of the 1-hectare site. It should be noted that most demonstrators did not submit the dig lists for Area A until the end of the demonstrations and therefore did not receive ground truth information to aid them in the analysis of Areas B and C.

(5) Items j through n were calculated from the prioritized dig lists as follows: Maximum achievable P(det) for each area was calculated as the number of items in the entire list that correspond to emplaced UXO targets (even though they may have been mis-discriminated as clutter) divided by the actual number of UXO targets emplaced in that site. Similarly, maximum achievable P(disc) was calculated as the number of anomalies in the dig list that were correctly classified as UXO divided by the total number of emplaced UXO targets. In addition, the single point P(disc) was determined by calculating the number of actual UXO targets that correctly classified as UXO and included in the list at or above the specified dig point. The operating (single point) FAR was calculated as the number of items per surveyed area included above the dig threshold and which did not correspond to emplaced target or emplaced clutter locations. FAR is therefore a measure of the false positives due to natural geologic/environmental factors and any pre-existing metal objects. In addition, Pfp was calculated as the ratio of the number of clutter items declared as UXO to the number of clutter items emplaced. Total FARs (item n) were computed by combining both false alarm sources that make up items

l and m. The government developed ROC-like curves developed by varying the dig threshold until the maximum P(det) and P(disc) were reached, and by plotting P(det), P(disc) as the ordinate and Pfp and FAR as the abscissa at each increment. ROC-like curves of P(det) and P(disc) versus total FAR were also developed by using the specified "signal strength" parameter as the thresholding variable. Performance comparisons between systems include using the ROC-like curves to determine Pfp and FAR at the P(det) required for Kaho'olawe Tier II clearance (85%).

(6) After each demonstrator had submitted the dig sheets described above, a total of three dig lists, he or she had an opportunity to re-analyze the data from the 1-hectare site and the six 30 m by 30 m grids to develop two additional prioritized dig sheets that take into account only targets larger than 40 mm projectiles. All dig sheets were to be submitted to the government representative within 3 days after completing the field demonstration and prior to departing Hawaii. After all field demonstrations were completed, the ESTCP Program Office provided each demonstrator with the complete ground truth for all the test areas at JPG and Kaho'olawe. The demonstrators were required to re-analyze his/her results, identify problems and potential improvements, and submit self-evaluation draft reports to the ESTCP Program Office.

(7) Item o, Operational costs were estimated using the cost factors and procedures described in Section 3.2.9.

3.2.7 Performance Assessment

In accordance with the Demonstration Plan, each demonstrator was responsible for determining the best method of employing his system to: (a) ensure full coverage of each demonstration area, (b) collect high-quality sensor data to support detection and discrimination requirements, (c) achieve high production rates, and (d) minimize man-hour requirements and costs. All demonstrators were able to complete the field surveys within the allotted time periods. There was a wide range in the demonstrators' survey data quality, data density, quality of analysis, and compliance with the data submission requirements specified in the Demonstration Plan (Reference 8). For example, a number of demonstrators failed to include required dig list information such as recommended stop dig points, appropriate confidence levels, and signal strength levels, and most demonstrators failed to re-analyze their data and prepare dig lists that excluded the small 20 mm and 40 mm targets. This lack of adherence to the requirements of the Demonstration Plan made the interpretation of results and adequate across-demonstrator performance comparisons very difficult. This Section presents a summary of the data the demonstrators submitted and the government's assessment of their performance.

It should be noted that because all the demonstrators submitted very high numbers of false alarms, the government was not able to fully investigate the sources of all of them. Nevertheless, during April 2002, NAVEODTECHDIV personnel conducted extensive surveys and excavation activities in the calibration and demonstration areas to verify the emplaced target locations and to attempt to identify the sources of many of the false alarms. Information from these post-demonstration activities was incorporated into the ground-truth data used to evaluate the demonstrators' performance.

3.2.7.1 Detection Results

Table 6 summarizes the number of UXO targets each demonstrator detected within a 1.0 m circular error and shows maximum achievable P(det). The P(det) is calculated as the number of declared items that correspond to emplaced UXO targets (even though they may have been misclassified as clutter) divided by the actual number of UXO targets emplaced in the demonstration site. Table 7 summarizes the detection results achieved when the small UXO targets (20 mm and 40 mm projectiles) were excluded from the evaluation. Tables 8 and 9 summarize the detection results within 0.5 m circular error.

Table 6. P(det) by Area Within 1.0 m.

		Within 1.0 m			
		Area A	**Area B**	**Area C**	**Total**
	Number of Actual Targets	24	81	34	139
NAEVA	Targets Detected	18	45	20	83
	P(det)	0.750	0.556	0.588	0.597
GTL	Targets Detected	18	41	18	77
	P(det)	0.750	0.506	0.529	0.554
Geophex	Targets Detected	20	46	15	81
	P(det)	0.833	0.568	0.441	0.583
NRL	Targets Detected	10	23	8	37
	P(det)	0.417	0.284	0.235	0.266
NRL without 20 and 40 mm	Targets Detected	11	33	12	56
	P(det)	0.458	0.407	0.353	0.403
Parsons EM61 EM-and-Flag	Targets Detected	18	50	22	90
	P(det)	0.750	0.617	0.647	0.647
Parsons EM61 Digital	Targets Detected	18	49	23	90
	P(det)	0.750	0.605	0.676	0.647
Parsons TM-5EMU	Targets Detected	7	22	4	33
	P(det)	0.292	0.272	0.118	0.237

These results indicate that none of the demonstrators were able to achieve the Kaho'olawe Tier II clearance requirements of 0.85 P(det) with 0.5 m location accuracy at any of the three demonstration areas. Only when the requirements were relaxed by expanding the allowable position error to 1.0 m and also deleting the smaller UXO targets, did any of the demonstrators meet the P(det) requirements. Even then, acceptable P(det) levels were only obtained in Area A which had considerably lower levels of geologic noise and metallic clutter than the other two areas.

Table 7. P(det) by Area Within 1.0 m and Without 20 and 40 mm

		Within 1.0 m			
		Area A	Area B	Area C	Total
	Number of Actual Targets without 20 and 40 mm	19	55	28	102
NAEVA	Targets Detected	16	38	19	73
	P(det)	0.842	0.691	0.679	0.716
GTL	Targets Detected	17	31	15	63
	P(det)	0.895	0.564	0.536	0.618
Geophex	Targets Detected	17	39	14	70
	P(det)	0.895	0.709	0.500	0.686
NRL	Targets Detected	9	22	8	37
	P(det)	0.474	0.400	0.286	0.363
NRL without 20/40 mm	Targets Detected	10	32	11	53
	P(det)	0.526	0.582	0.393	0.520
Parsons EM61 EM-and-Flag	Targets Detected	17	42	21	80
	P(det)	0.895	0.764	0.750	0.784
Parsons EM61 Digital	Targets Detected	16	39	22	77
	P(det)	0.842	0.709	0.786	0.755
Parsons TM-5EMU	Targets Detected	7	19	4	30
	P(det)	0.368	0.345	0.143	0.294

Table 8. P(det) by Area Within 0.5 m.

		Within 0.5 m			
		Area A	Area B	Area C	Total
	Number of Actual Targets	24	81	34	139
NAEVA	Targets Detected	13	33	10	56
	P(det)	0.542	0.407	0.294	0.403
GTL	Targets Detected	10	30	3	43
	P(det)	0.417	0.370	0.088	0.309
Geophex	Targets Detected	16	33	10	59
	P(det)	0.667	0.407	0.294	0.424
NRL	Targets Detected	8	17	6	23
	P(det)	0.333	0.210	0.176	0.165
NRL without 20/40 mm	Targets Detected	9	19	8	36
	P(det)	0.375	0.235	0.235	0.259
Parsons EM61 EM-and-Flag	Targets Detected	12	33	15	60
	P(det)	0.500	0.407	0.441	0.432
Parsons EM61 Digital	Targets Detected	12	25	15	52
	P(det)	0.500	0.309	0.441	0.374
Parsons TM-5EMU	Targets Detected	6	13	1	20
	P(det)	0.250	0.160	0.029	0.144

Table 9. P(det) by Area Within 0.5 m and Without 20 and 40 mm.

		Within 0.5 m			
		Area A	Area B	Area C	Total
	Number of Actual Targets without 20 and 40 mm	19	55	28	102
NAEVA	Targets Detected	11	26	9	46
	P(det)	0.579	0.473	0.321	0.451
GTL	Targets Detected	9	24	3	36
	P(det)	0.474	0.436	0.107	0.353
Geophex	Targets Detected	13	28	9	50
	P(det)	0.684	0.509	0.321	0.490
NRL	Targets Detected	7	17	6	23
	P(det)	0.368	0.309	0.214	0.225
NRL without 20 and 40 mm	Targets Detected	8	19	8	35
	P(det)	0.421	0.345	0.286	0.343
Parsons EM61 EM-and-Flag	Targets Detected	11	26	15	52
	P(det)	0.579	0.473	0.536	0.510
Parsons EM61 Digital	Targets Detected	10	24	15	49
	P(det)	0.526	0.436	0.536	0.480
Parsons TM-5EMU	Targets Detected	6	12	1	19
	P(det)	0.316	0.218	0.036	0.186

Tables 10 and 11 summarize the performance of the demonstrators over the entire demonstration site within the QA range for 0.5 m and 1.0 m circular halos, respectively. In these tables, the ordnance items left in the ground represent the items declared by the demonstrator as clutter with high confidence that are actually ground truth emplaced ordnance plus the number of undetected ordnance items. The variable, correct discrimination, is the total of the correctly identified ordnance items. The False Alarm number is the total of the other detections plus the total of the ground truth fragment matches minus the number of objects identified as clutter with high confidence.

Table 10. Summary of Discrimination Performance Within 0.5 m Circular Halo.

	Ordnance Left in the Ground		Correct Discrimination		False Alarm Number
	Within 0.5 m		Within 0.5 m		Within 0.5 m
	Num	Percent	Num	Percent	
NAEVA	88	63.31%	51	36.69%	624
GTL	96	69.06%	43	30.94%	1,283
Geophex	80	57.55%	59	42.45%	772
Geophex without 20 and 40 mm	80	57.55%	59	42.45%	772
NRL	108	77.70%	31	22.30%	342
NRL without 20 and 40 mm	102	73.38%	37	26.62%	602
Parsons (EM61) EM-and-Flag	79	56.83%	60	43.17%	872
Parsons EM61 Digital	87	62.59%	52	37.41%	1,405
Parsons TM-5 EMU	119	85.61%	20	14.39%	172

Table 11. Summary of Discrimination Performance Within 1.0 m Circular Halo.

	Ordnance Left in the Ground		Correct Discrimination		False Alarm Number
	Within 1.0 m		Within 1.0 m		Within 1.0 m
	Num	Percent	Num	Percent	
NAEVA	62	44.60%	77	64.03%	842
GTL	62	45.32%	77	55.40%	1,249
Geophex	58	41.73%	81	58.27%	750
Geophex w/out 20 and 40 mm	58	41.73%	81	58.27%	750
NRL	98	70.50%	41	29.50%	333
NRL without 20 and 40 mm	83	59.71%	56	40.29%	584
Parsons (EM61) EM-and-Flag	49	35.25%	90	64.75%	854
Parsons EM61 Digital	49	35.25%	90	64.75%	1,530
Parsons TM-5 EMU	106	76.26%	33	23.74%	167

3.2.7.2 ROC-Based Performance Assessment

Figures 37 through 51 show the results obtained when the location accuracy is set to 0.5 m. The analysis is also performed with the location accuracy set to 1.0 m and is in Appendix C of Reference 11. The performance of all of the demonstrators in all cases fell below the Kaho'olawe Tier II clearance requirements. In Figure 37, Geophex demonstrated significantly better performance that the other demonstrators, and in Figures 39, 45, 48, and 49, GTL demonstrated significantly poorer performance.

Figures 37, 38, and 39 show the demonstrators' detection performance for Areas A, B, and C, respectively. The distance threshold for scoring a detection is set to 0.5 m for these sets of plots. Figure 37 shows that in Area A all the systems operated along very similar ROC curves, with the major difference being their selection of the "stop dig" or endpoint threshold. In the cases of NRL and Parsons TM-5 EMU, the endpoint thresholds were set so high that the operating (and maximum achievable) P(det) was much lower than those of the other demonstrators. Geophex achieved the best performance for this area, but it is apparent from the steep slope of the ROC curve that a lower endpoint threshold would probably have resulted in increased P(det) with a relatively small increase in Total FAR.

The placement of the "stop dig" point (shown as the SD triangle in each of these figures) indicates the demonstrators' general lack of confidence in their discrimination capability. With the exception of NAEVA, all demonstrators placed their "stop dig" point at the end of the dig list. NAEVA's attempt at discrimination resulted in an operating P(det) that was 10% lower than the maximum achievable in Area A.

Figure 39 shows that all demonstrators, with the exception of GTL, achieved very similar ROC-based performance. GTL demonstrated significantly poorer detection and Total FAR performance than the others. Overall, all of the systems demonstrated poorer performance in Area B than Area A, confirming previous reports that Area B contained significantly higher levels of geologic anomalies and metallic clutter. As in Area A, all demonstrators except NAEVA were extremely conservative in their "stop dig" point selection.

Figures 40 through 42 show the results of re-analyzing the data to include only targets larger than 40 mm projectiles. Only NRL and Geophex performed this analysis and resubmitted revised dig lists.

The small number of submissions and the lack of this type of data from Parsons precludes any comparisons across systems or with a baseline. In order to overcome this deficiency, the demonstrators' dig lists that contained all targets (including 20 mm and 40 mm) were evaluated with the 20 and 40 mm targets removed from the ground truth. No detection, false alarm, or missed target was assigned to any declaration within 1.0 m of these small emplaced targets.

Figures 43 through 45 show the results of this evaluation. These ROC curves show a slight improvement in performance for all demonstrators, but no significant change in ROC curve shape or in the relative performance between demonstrators. It can be concluded that all systems have similar capability (or lack thereof) to detect the smaller targets in the Kaho'olawe environment.

Figures 46 through 48 show the P(det) performance as a function of Pfp for all demonstrators, where Pfp is computed as the ratio of the number of emplaced clutter items included on the dig list to the number of clutter items emplaced. This metric attempted to separate the effects of the geology and unknown metal clutter from those known, emplaced clutter items. These ROC curves show very small differences between demonstrators, but most importantly, the almost consistently flat, diagonal shape of the curves indicates that none of the systems demonstrated a capability to discriminate emplaced UXO from emplaced metallic clutter in this environment.

Figures 49 through 51 show the systems' ability to discriminate ordnance from nonordnance (both geologic and metallic) based solely on signal strength. The ROC curves for all demonstrators tend to fall within a narrow band and do not support any conclusions regarding the various systems' discrimination capability.

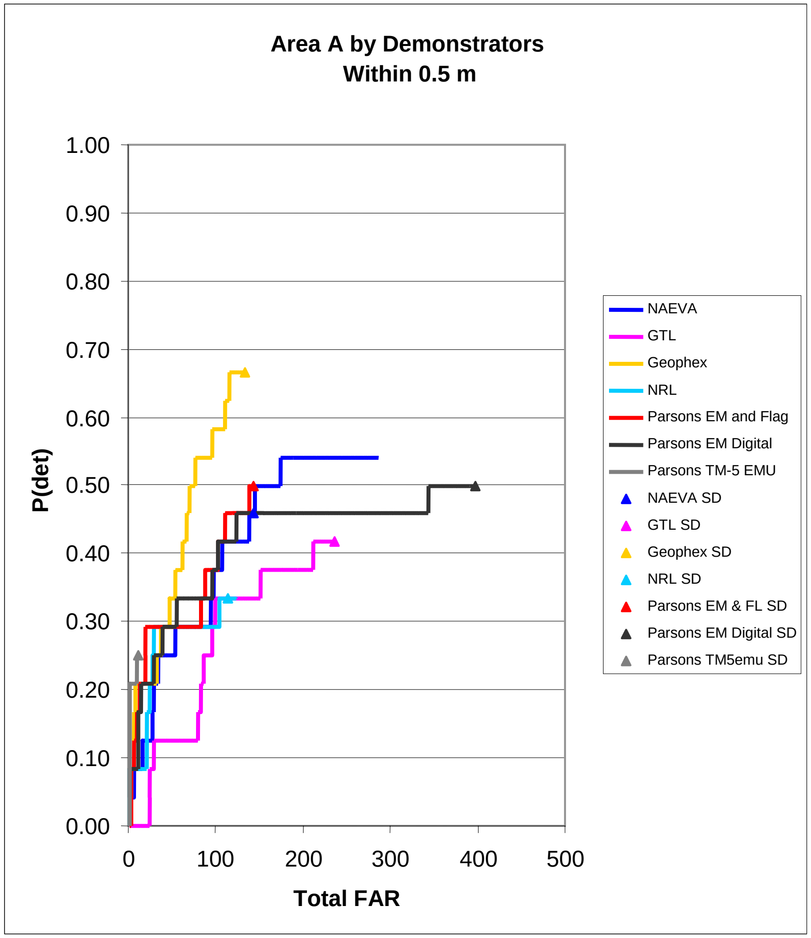

Figure 37. Area A—P(det) Versus Total FAR Within 0.5 m for All Demonstrators.

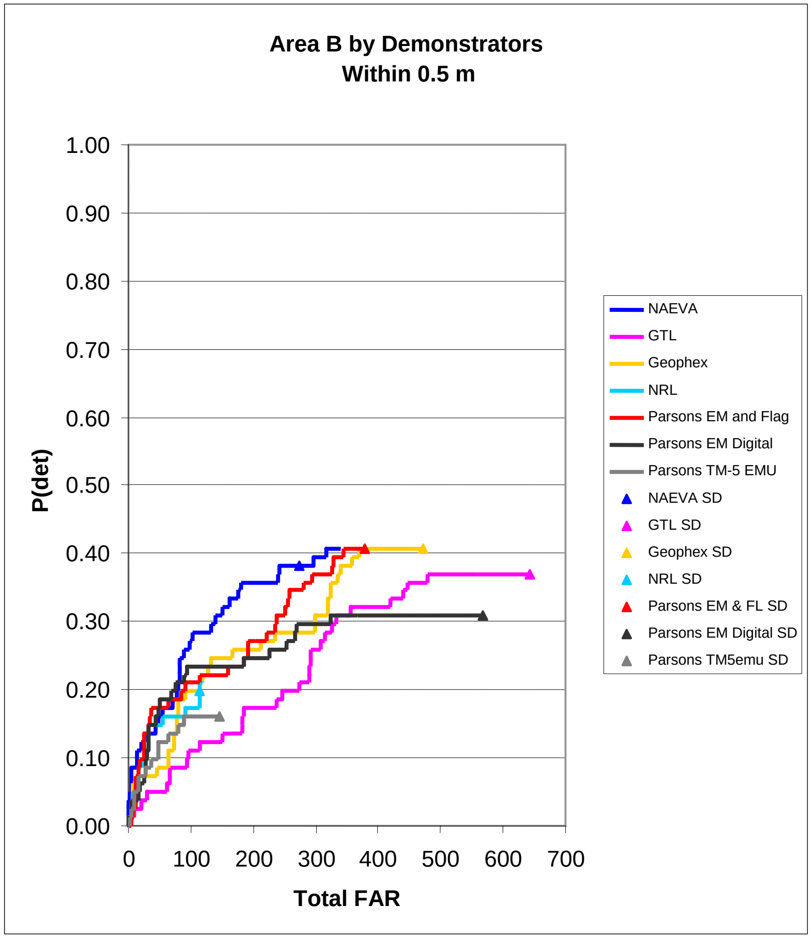

Figure 38. Area B—P(det) Versus Total FAR Within 0.5 m for All Demonstrators.

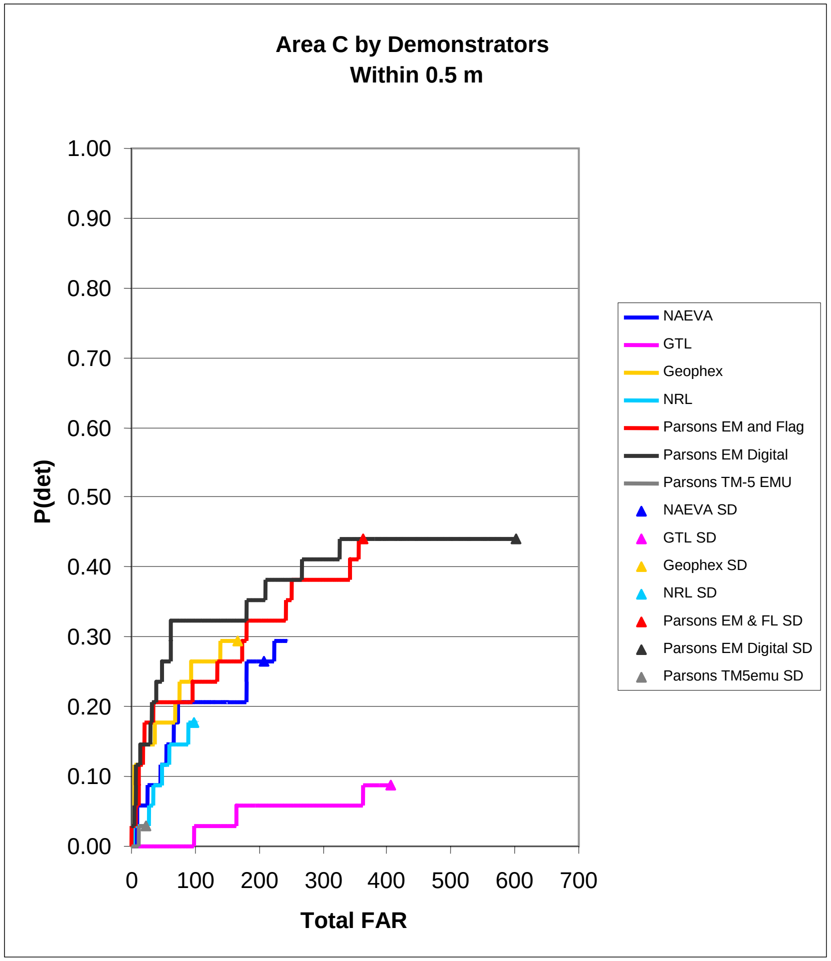

Figure 39. Area C—P(det) Versus Total FAR Within 0.5 m for All Demonstrators.

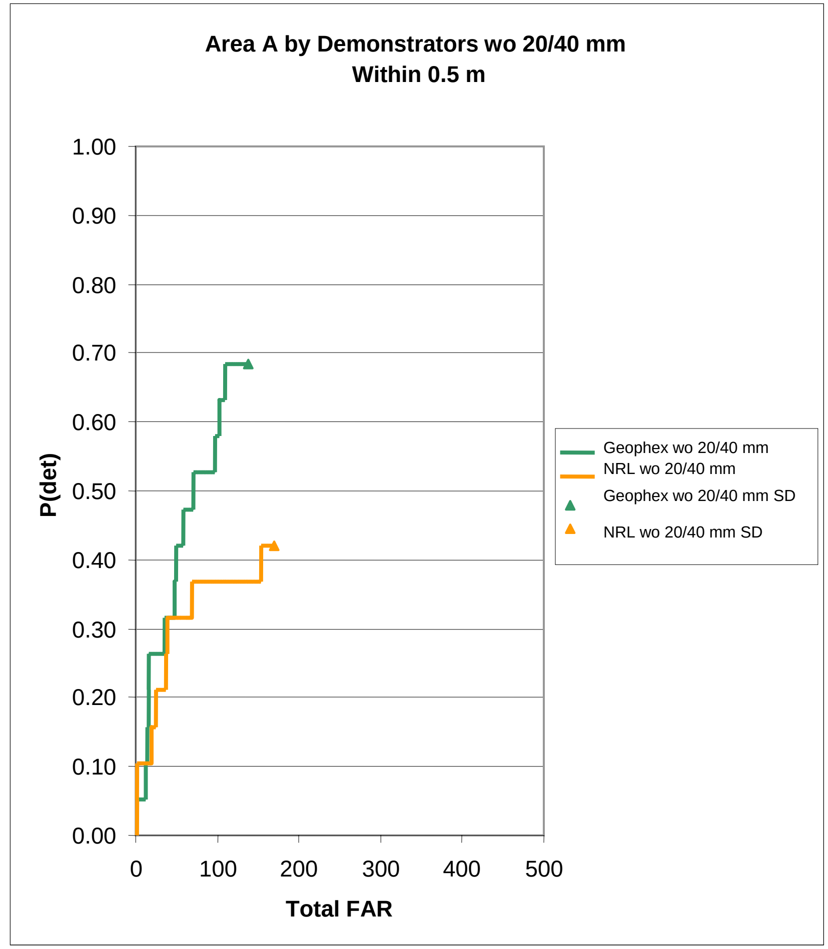

Figure 40. Area A—P(det) Versus Total FAR Within 0.5 m for Demonstrators Without 20/40 mm.

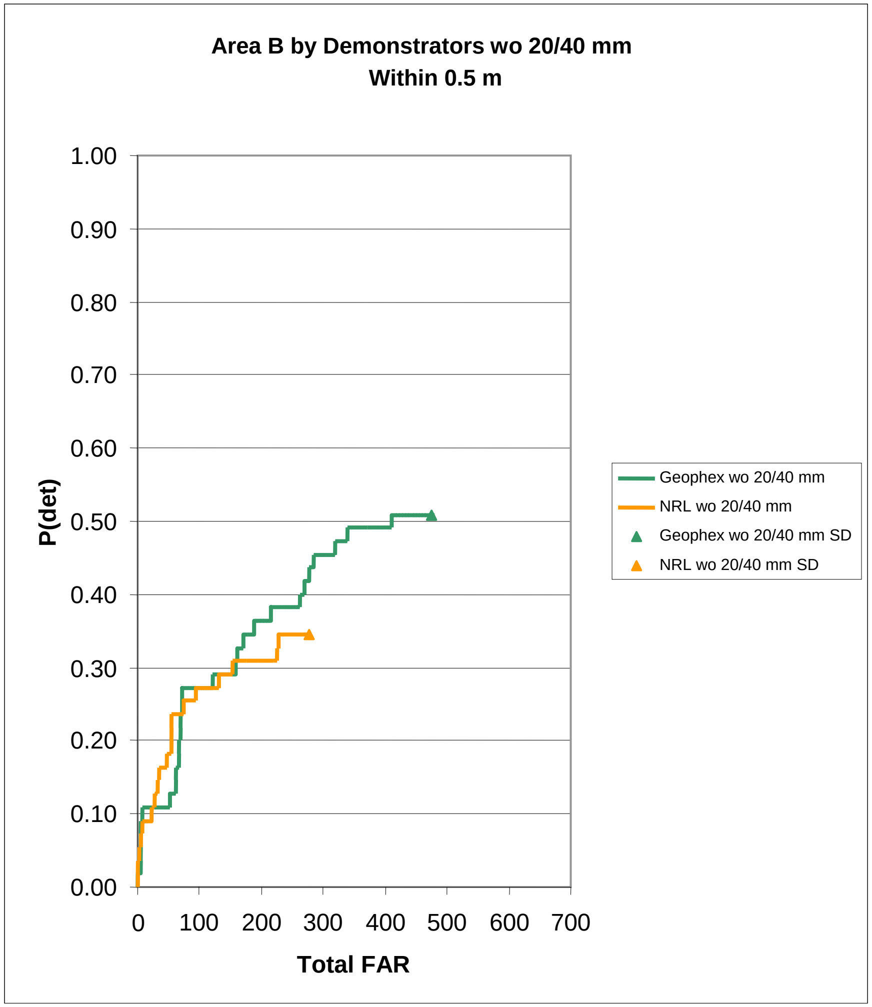

Figure 41. Area B—P(det) Versus Total FAR Within 0.5 m for Demonstrators Without 20/40 mm.

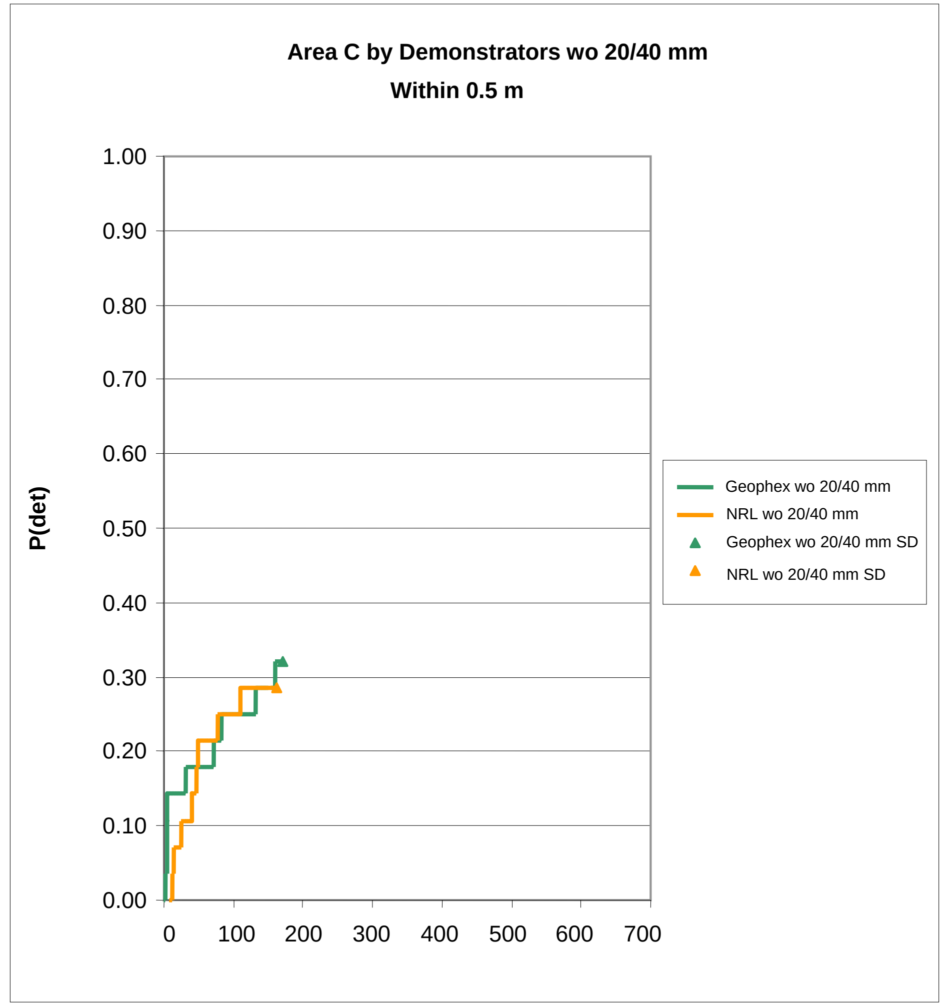

Figure 42. Area C—P(det) Versus Total FAR Within 0.5 m for Demonstrators Without 20/40 mm.

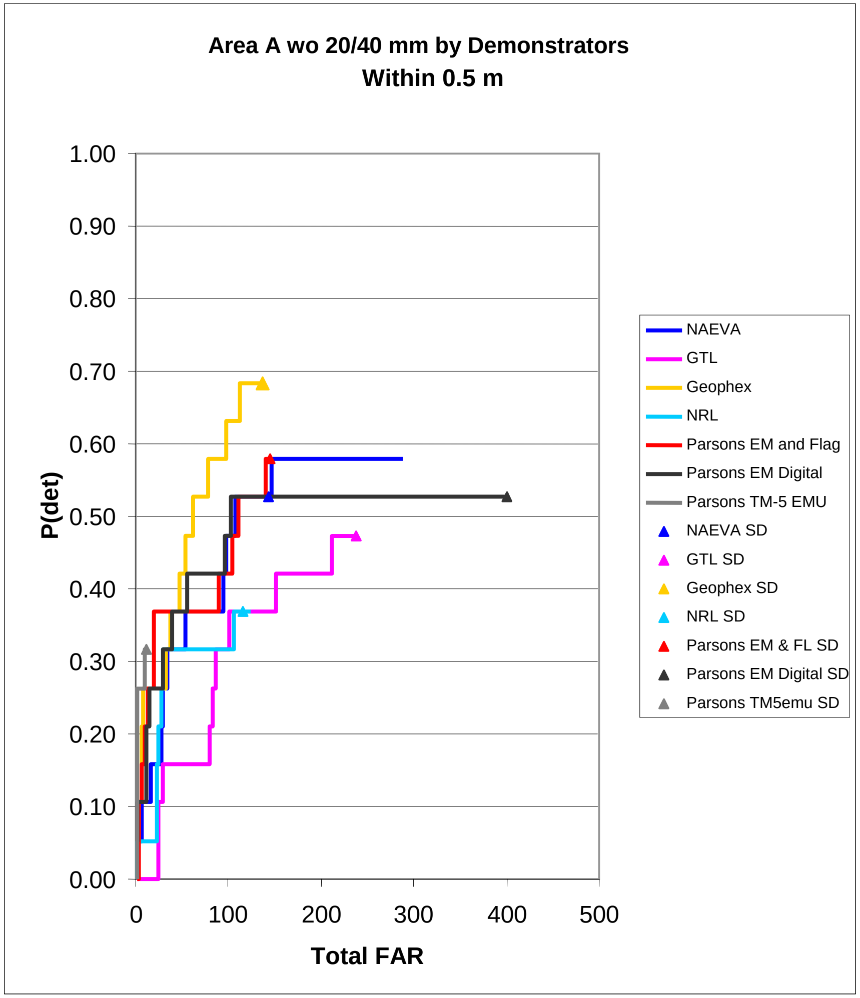

Figure 43. Area A Without 20 and 40 mm—P(det) Versus Total FAR Within 0.5 m for All Demonstrators.

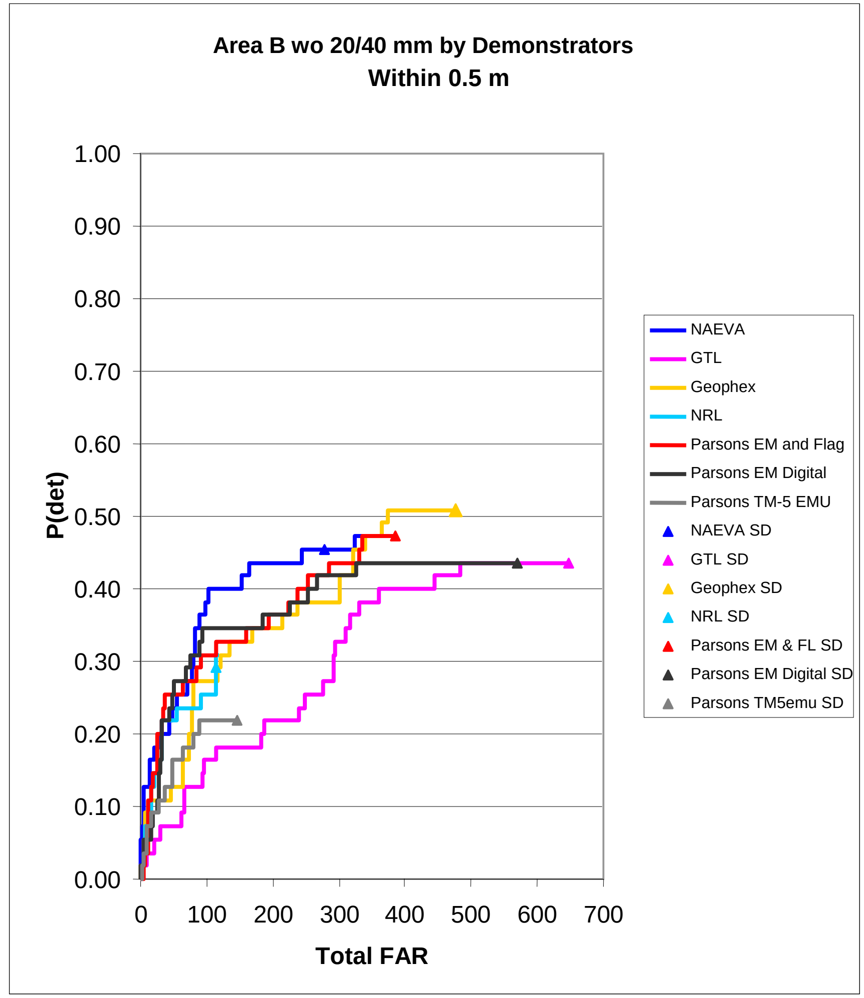

Figure 44. Area B Without 20 and 40 mm—P(det) Versus Total FAR Within 0.5 m for All Demonstrators.

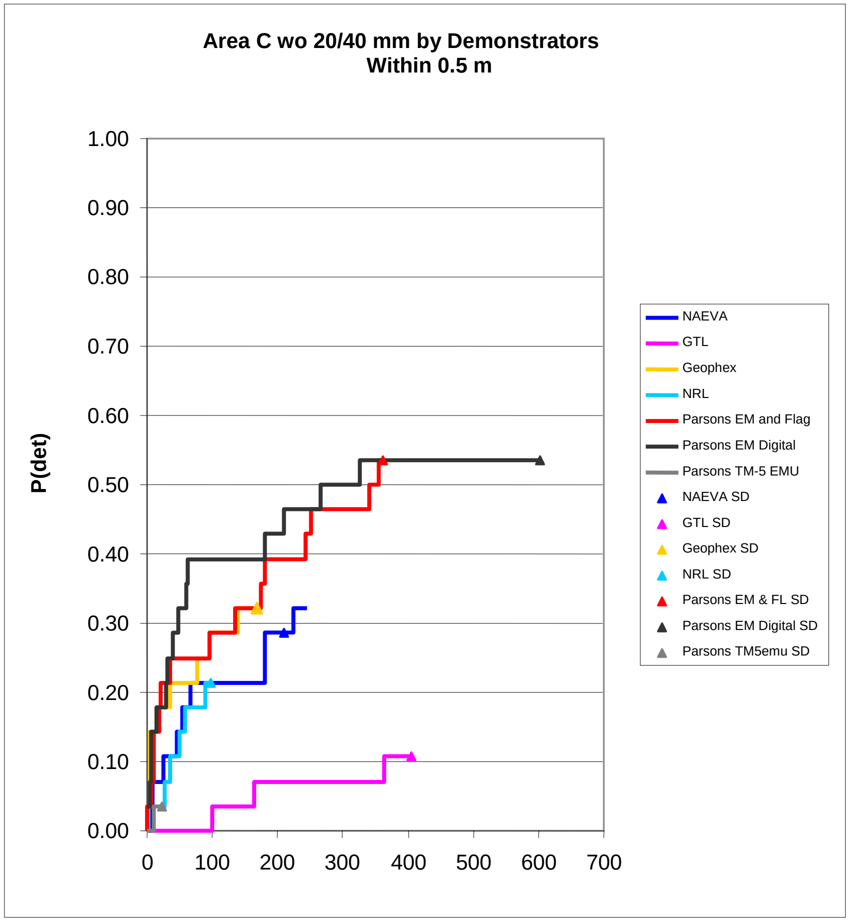

Figure 45. Area C Without 20 and 40 mm—P(det) Versus Total FAR Within 0.5 m for All Demonstrators.

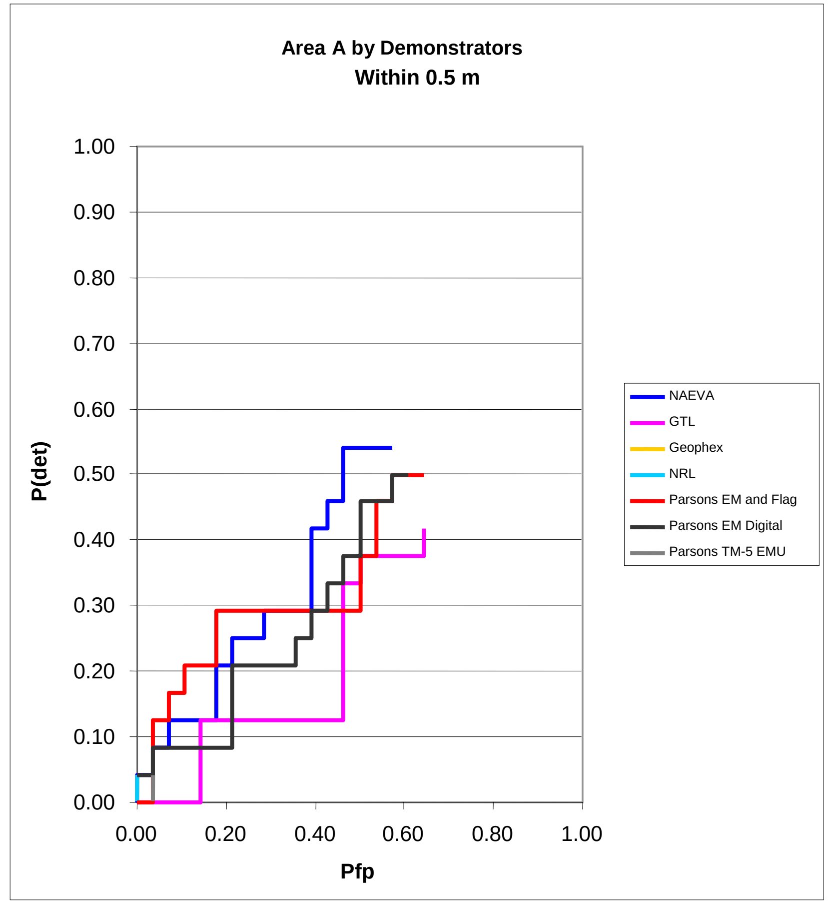

Figure 46. Area A—P(det) Versus Pfp Within 0.5 m for All Demonstrators.

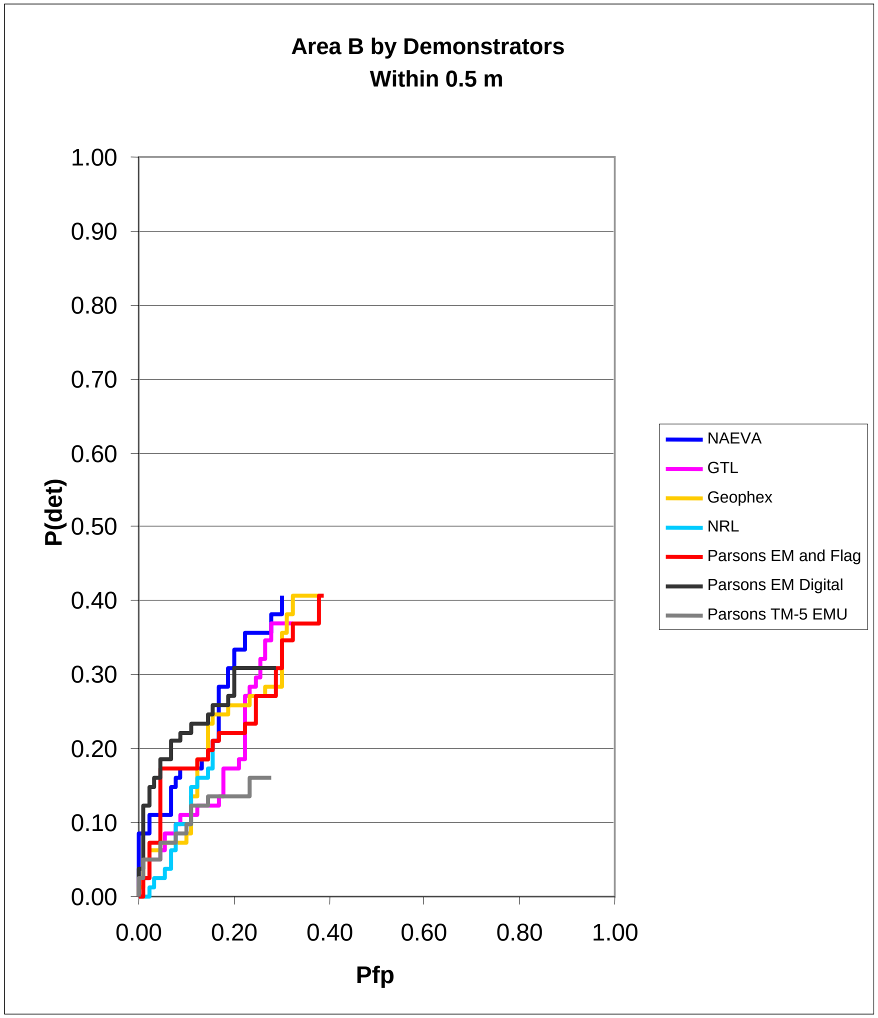

Figure 47. Area B—P(det) Versus Pfp Within 0.5 m for All Demonstrators.

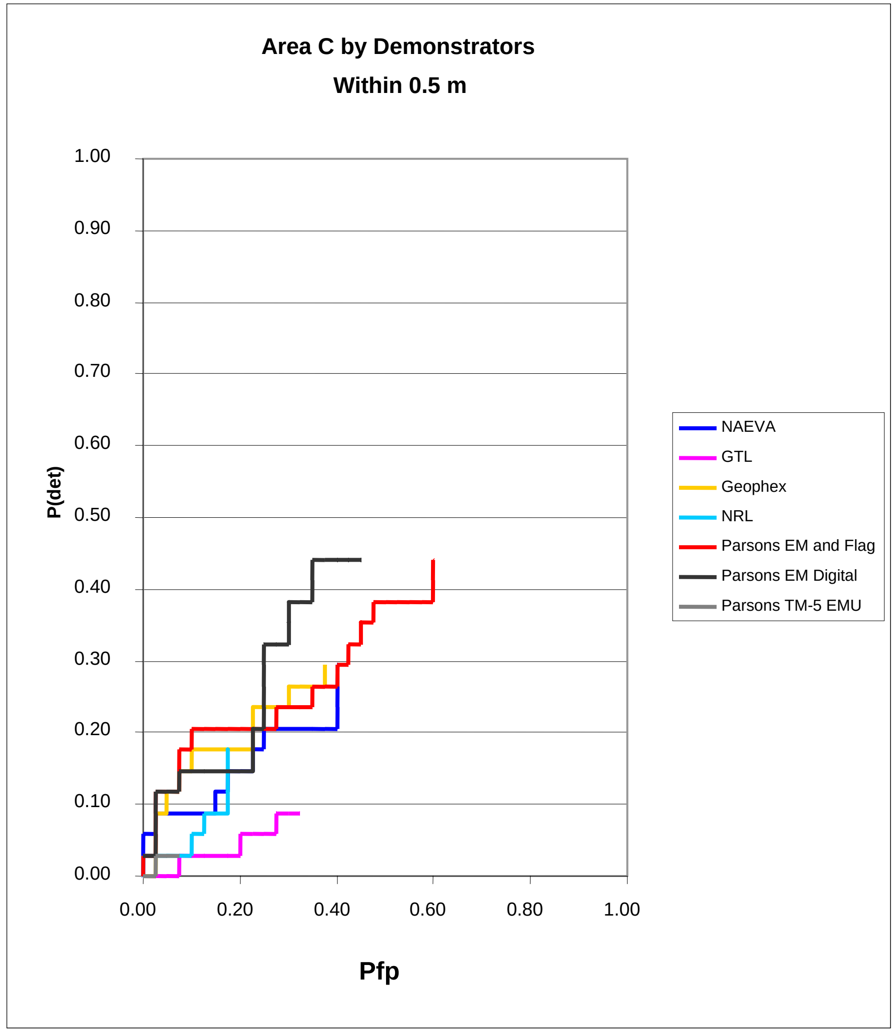

Figure 48. Area C—P(det) Versus Pfp Within 0.5 m for All Demonstrators.

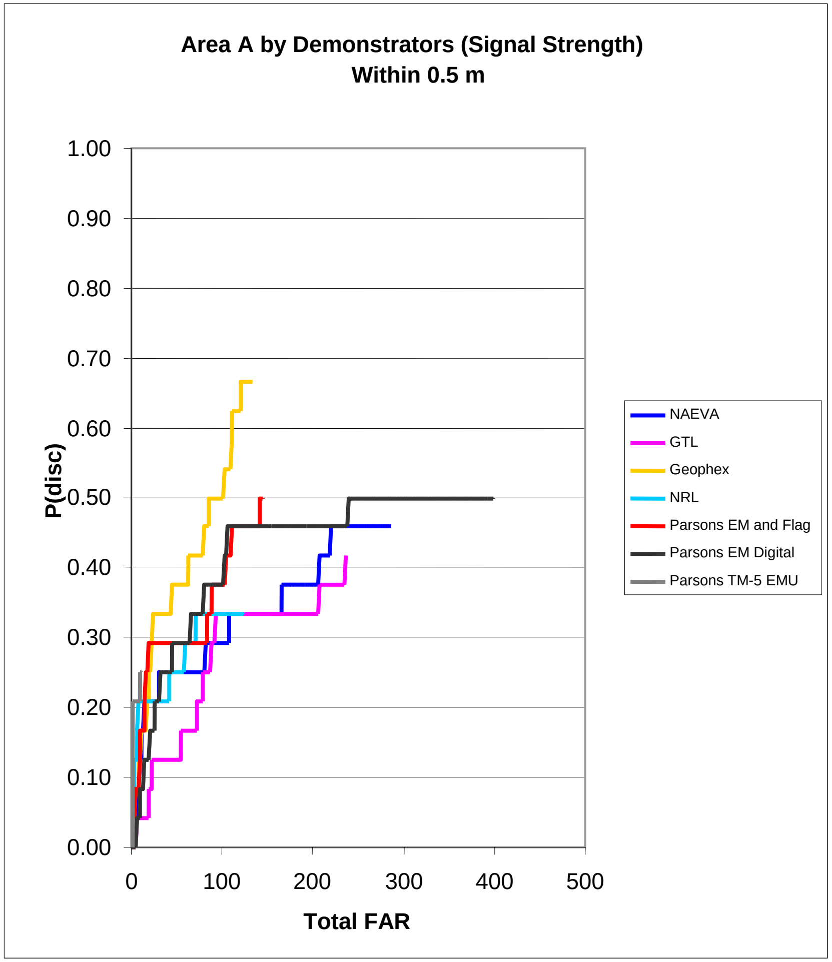

Figure 49. Area A—Pdisc (Prioritized by Signal Strength) Versus Total FAR Within 0.5 m for All Demonstrators.

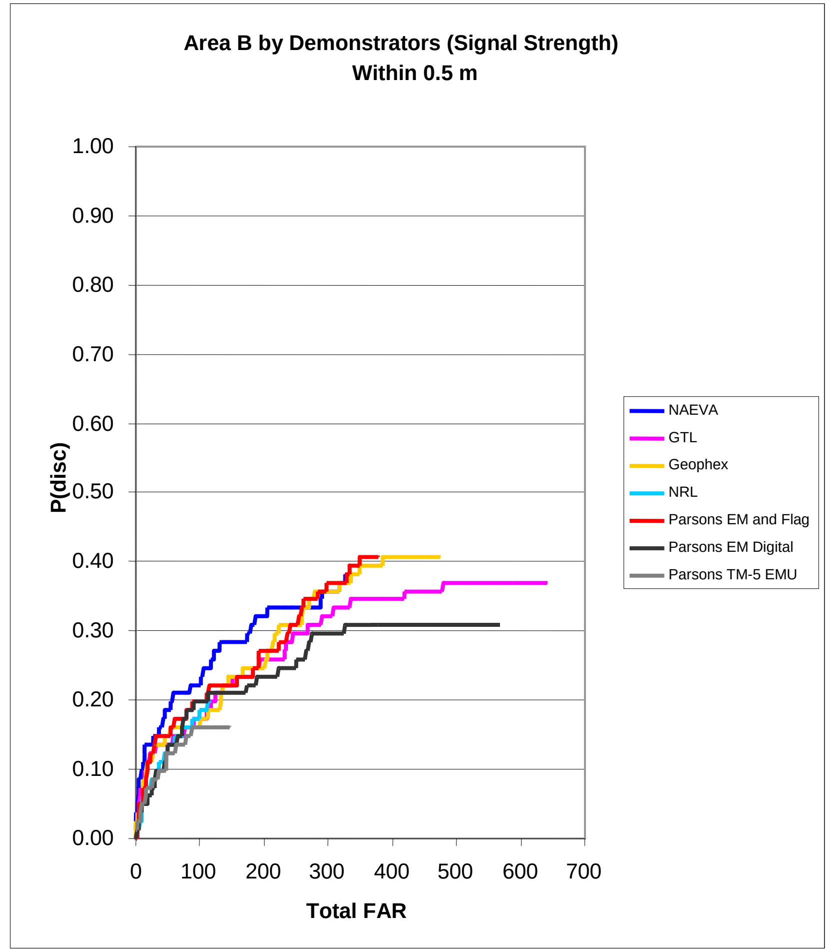

Figure 50. Area B—Pdisc (Prioritized by Signal Strength) Verus Total FAR Within 0.5 m for All Demonstrators.

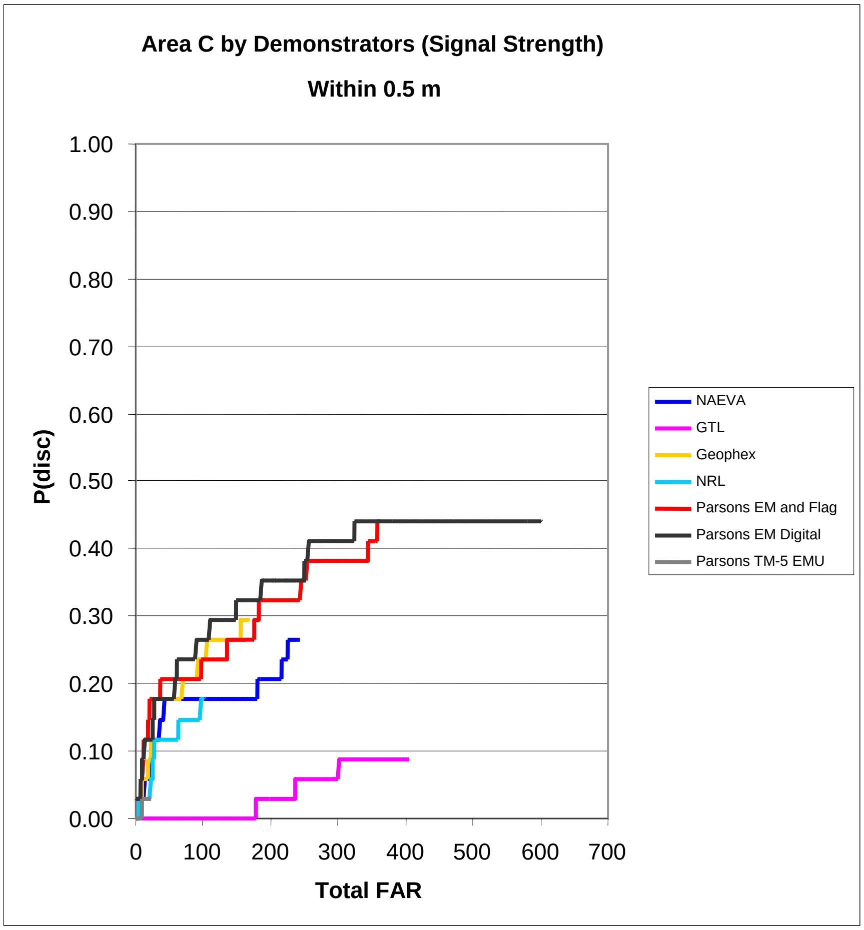

Figure 51. Area C—Pdisc (Prioritized by Signal Strength) Versus Total FAR Within 0.5 m for All Demonstrators.

3.2.8 Technologies Comparison

The detection performance of all demonstrated systems was considerably lower than expected and significantly lower than those demonstrated during prior field demonstrations such as JPG Phases II-IV and during the first phase of this project. None of the systems demonstrated any ability to discriminate ordnance from metallic clutter, much less to identify ordnance by size or type. It was not possible to evaluate the systems' ability to discriminate ordnance from geologic anomalies because the demonstration area was so cluttered with unknown metallic objects that the effects of geology could not be reliably separated. Some obvious reasons for the decreased overall performance include the facts that, unlike the fairly benign, low noise environment of test sites such as JPG, Kaho'olawe presents an extreme in clutter density, geologic noise, and difficult operating environment.

Overall, the best performances were achieved in Area A where NAEVA, GTL, Geophex, Parsons EM-and-flag, and Parsons EM61 digital had a maximum achievable P(det) of at least 50% when including detections within 0.5 m. Also, the Total FAR was the lowest for all demonstrators in Area A. However, none of the demonstrators reached the P(det) required for Kaho'olawe Tier II clearance (85%). The P(det) in Areas B and C were significantly lower for all demonstrators and their Total FAR was higher. Most significantly, none of the advanced EMI systems demonstrated significant improved capability over the baseline EM-61 system operated in the EM-and-Flag mode.

3.2.9 Cost Assessment

Labor costs associated with each field task were computed by applying the cost factors described in the Demonstration Work Plan and detailed in Table 12. It should also be noted that since all the demonstrated systems were man-portable or hand held, with similar support equipment and capital cost requirements, it was assumed that mobilization/demobilization and life-cycle costs would be equal and could be omitted from this relative cost performance evaluation.

Analysis of this table indicates that the field labor costs of most of the demonstrators fall within a fairly narrow range, with the two exceptions being the Parsons TM-5 EMU, which demonstrated significantly lower costs, and NAEVA, which achieved significantly higher costs. These differences may be attributable to the fact that the TM-5EMU survey conducted by Parsons did not appear to adequately cover the required areas (based on their very low number of detected anomalies and on the fact that GTL, using the same sensor, required significantly more time to complete the surveys) and to the fact that NAEVA approached this field demonstration effort from a more scientific perspective. (One of their primary objectives was to collect very high quality field data to support their ongoing algorithm development efforts.) With the exception of NAEVA, all the other advanced EMI systems compare favorably against the EM-61 EM-and-Flag baseline technology.

Table 13 summarizes the operational costs of the demonstrator systems after the cost penalties described in Chapter 4 were applied. These penalties consisted of $200 for each false alarm (clutter item selected for digging by the demonstrator), and the cost of a complete resurvey for one or more UXO targets missed (not included in the dig list) or erroneously declared as clutter with high confidence. This table highlights the fact that false alarms have (by a large margin) the greatest impact on the cost performance of each system. Table 13 indicates that all seven demonstrators were penalized with the cost of a resurvey at each of the three demonstration areas because their dig lists indicate that UXO would have had been left in the ground. It is difficult to draw conclusions regarding the performance of each system based solely on these results. For example, one may conclude that the Parsons TM-5 EMU system is superior because of its significantly lower costs, but analysis of their detection performance would show that it failed to detect a very large percentage of the UXO present. Since the cost of false alarms dominates this type of analysis in very highly cluttered areas such as Kaho’olawe, any system could achieve low costs by operating on a very low point on the ROC curve. Thus, any cost comparisons between systems have to be correlated with their respective ROC curve performance in order to reach reasonable conclusions. Assuming that the systems are operated at reasonable points of the ROC curve, Table 12 would indicate that Parsons EM-61 EM-and-Flag and NAEVA are the most cost-effective performers. Another factor to be considered in future evaluations is the fact that the $200 cost per false alarm is excessive for sites such as Kaho’olawe where the high density and fairly shallow depths of most metallic clutter from munitions fragments allows for rapid removal with minimal manpower requirements.

Table 12. Breakdown of Field Costs.

Demonstrator	Number of Operators	Categories	Hourly Rate	Time (hrs, min)	Cost of Job
NAEVA	1	Supervisor	$95.00	35:45	$3,396.25
	3	Logistic/Field Setup	28.50	15:30	441.75
	3	Logistic/Field Survey	28.50	83:00	2,365.50
	3	Logistic/Field Downtime	28.50	1:30	42.75
	3	Logistic/Field Resurvey	28.50	6:30	185.25
	4	Total			**$6,431.50**
GTL	1	Supervisor	$95.00	21:40	$2,058.33
	1	Logistic/Field Setup	28.50	3:45	106.88
	1	Logistic/Field Survey	28.50	13:43	390.92
	1	Logistic/Field Downtime	28.50	3:34	101.65
	1	Logistic/Field Resurvey	28.50		
	2	**Total**			**$2,657.78**
Geophex	1	Supervisor	$95.00	24:30	$2,327.50
	2	Logistic/Field Setup	28.50	10:30	299.25
	2	Logistic/Field Survey	28.50	24:25	688.75
	2	Logistic/Field Downtime	28.50	2:20	66.50
	2	Logistic/Field Resurvey	28.50	2:55	83.12
	3	**Total**			**$3,465.12**
NRL	1	Supervisor	$95.00	20:02	$1,903.17
	3	Logistic/Field Setup	28.50	9:00	256.50
	3	Logistic/Field Survey	28.50	34:06	977.55
	3	Logistic/Field Downtime	28.50	19:00	541.50
	3	Logistic/Field Resurvey	28.50		
	4	Total			**$3,678.72**
Parsons (EM61) EM-and-Flag	1	Supervisor	$95.00	28:01	$2,661.58
	2	Logistic/Field Setup	28.50	18:57	540.08
	2	Logistic/Field Survey	28.50	58:30	1,667.25
	2	Logistic/Field Downtime	28.50	17:00	484.50
	2	Logistic/Field Resurvey	28.50		
	3	**Total**			**$5,353.41**
Parsons EM61 Digital	1	Supervisor	$95.00	17:55	$1,702.08
	1	Logistic/Field Setup	28.50	16:44	476.60
	1	Logistic/Field Survey	28.50	15:04	446.50
	1	Logistic/Field Downtime	28.50	7:16	207.10
	1	Logistic/Field Resurvey	28.50	2:50	80.75
	2	**Total**			**$2,913.33**
Parsons TM-5 EMU	1	Supervisor	$95.00	6:56	$ 658.67
	3	Logistic/Field Setup	28.50	12:10	345.75
	3	Logistic/Field Survey	28.50	5:00	142.50
	3	Logistic/Field Downtime	28.50	5:30	156.75
	3	Logistic/Field Resurvey	28.50		
	4	Total			**$1,304.67**

Table 13. Demonstrator Costs, Including Penalties for False Alarms and for Leaving UXO Targets in Ground.

	NAEVA	GTL	Geophex	NRL	Parsons (EM61) (EM-and-Flag)	Parsons TM-5 EMU	Parsons Digital
Cost to Survey	$2,366	$391	$689	$978	$1,667	$143	$447
Cost of Resurvey	2,366	391	689	978	1,667	143	447
Cost of False Alarms	128,000	260,600	157,800	123,600	176,200	43,600	319,400
Total Cost	$132,732	$261,382	$159,178	$125,556	$179,534	$43,886	$320,294

3.2.10 Lessons Learned

The most surprising conclusion from this demonstration was the fact that none of the advanced EMI systems demonstrated significant performance and/or cost improvements over the baseline technology consisting of a standard EM-61 system operated in an EM-and-Flag mode. The relatively good performance achieved by this system in this mode would indicate that there may be advantages to providing real-time feedback to the UXO survey crews so they may collect additional sensor data (in orthogonal directions) over suspected anomalies, rather than blindly surveying lanes with fixed lane widths and sampling rates. Such a survey method would also allow the crew to visually identify and mark surface anomalies that could otherwise be misinterpreted as UXO during the post-survey data analysis.

Another important lesson learned from this demonstration is the difficulty in setting up test sites and conducting demonstrations at live UXO sites that are in the process of being remediated. Even though the calibration and demonstration areas had been cleared numerous times, there were still excessive amounts of metallic clutter from unknown sources, and even live ordnance that remained in these areas. As a result, the accuracy of the ground truth available for such test sites is always in doubt. In addition, the presence of unknown metallic clutter prevented the evaluation of the advanced EMI systems' assumed capability to mitigate the effects of geologic noise. Finally, the safety and logistics problems associated with conducting technology demonstrations concurrent with actual UXO cleanup operations proved to be a very inefficient, costly, and time consuming process.

It can be concluded from these demonstrations that additional research, development, and demonstration work is needed in order to produce UXO technologies that meet reasonable detection, discrimination, and false alarm performance goals, especially in difficult sites such as those encountered at Kaho'olawe, Hawaii.

This page left blank intentionally.

4.0 REFERENCES

1. "Advanced UXO Detection/Discrimination Technology Demonstration, U.S. Army Jefferson Proving Ground, Madison, Indiana," Technology Demonstration Plan, Naval Explosive Ordnance Disposal Technology Division (NAVEODTECHDIV), 30 June 00.

2. "The Parsons-UXB Express," Volume 2, Issue 3, 16 Mar 00.

3. Naval Explosive Ordnance Disposal Technology Division. (2000). "Site preparation plan—Advanced UXO detection/discrimination technology demonstration, U.S. Army Jefferson Proving Ground, Madison, Indiana," Washington, DC.

4. Geophysical Technology Limited. (2002). "Requested Information Regarding the GTL TM-5 EMU," Commercial-In-Confidence correspondence to the U.S. Army Corps of Engineers.

5. Won, I.J., Keiswetter, K., and Novikova, E. (1998). "Electromagnetic induction spectroscopy," Journal of Environmental and Engineering Geophysics 3, (1), 27-40.

6. Geonics Limited. (2000). "EM61 Metal Detector for Unexploded Ordnance (UXO) Detection and Characterization," Ontario, Canada.

7. Naval Research Laboratory. (2001). "Man-portable adjuncts for the MTADS," Technical Report NRL/PU/6110--01-434, Washington D.C.

8. Naval Explosive Ordnance Disposal Technology Division (NAVEODTECHDIV). (2001). "Technology Demonstration Plan—Advanced UXO Discrimination/ Discrimination Technology Demonstration, Kaho'olawe Island, Hawaii"

9. NAVEODTECHDIV. (2001). "Site Preparation Plan—Advanced UXO Detection/Discrimination Technology Demonstration, Kaho'olawe Island, Hawaii."

10. Cespedes, E. R. (2001). "Advanced UXO Discrimination/Discrimination Technology Demonstration—U.S. Army Jefferson Proving Ground, Madison, Indiana," Technical Report ERDC/EL TR-01-20, U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi.

11. Cargile, D. M., Bennett, H. H., Goodson, R. A., DeMoss, T. A., and Cespedes, E. R. (2002). "Advanced UXO Detection/Discrimination Technology Demonstration—Kaho'olawe Island, Hawaii," Technical Report ERDC/EL TR-02-XX, U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi.

12. Fairbanks, J.R., Baseline Performance of Traditional and Emerging UXO Detection Technologies: UXO Forum 2004, St. Louis, Missouri, March 9-12.

APPENDIX A

POINTS OF CONTACT

Point of Contact	Organization Name and Address	Phone/Fax/E-mail
Mr. Hien Dinh	NAVEODTECHDIV 2008 Stump Neck Road Indian Head, MD 20640-5070	Phone: 301-744-6850, Ext. 267 Fax: 301-744-6947 Email: dinh@eodpoe2.navsea.navy.mil
Dr. Ernesto Cespedes	USAEWES 3909 Halls Ferry Road Vicksburg, MS 39180-6199	Phone: 601-634-2655 Fax: 601-634-2732 Email: cespede@wes.army.mil
Mr. George Robitaille	U.S. Army Environmental Center ATTN: SFIM-AEC-P2/ETD Bldg. E4430 Aberdeen Proving Ground, MD 21010-5401	Phone: 410-436-6850 Fax: 410-436-6836 Email: George.Robitaille@aec.apgea.army.mil
Mr. Ken Knouf	Jefferson Proving Ground, IN	Phone: 812-273-6075
Dr. Joe Robb	Jefferson Proving Ground, IN	Phone: 812-273-0783

ESTCP Program Office

901 North Stuart Street
Suite 303
Arlington, Virginia 22203

(703) 696-2117 (Phone)
(703) 696-2114 (Fax)

e-mail: estcp@estcp.org
www.estcp.org